

JUHA LAINE

Laatua ja tuotannollista tehokkuutta?

Taloustieteellinen tutkimus vanhusten laitoshoidosta



Juha Laine

Laatua ja tuotannollista tehokkuutta?
Taloustieteellinen tutkimus vanhusten laitoshoidosta

Stakes
Sosiaali- ja terveysalan
tutkimus- ja kehittämiskeskus

Tutkimuksia 151

Toimitusneuvosto:

Matti Heikkilä, puheenjohtaja

Mika Gissler

Irma Kiikkala

Matti Kortteinen

Harri Sintonen

Sirpa Taskinen

Matti Virtanen

© Kirjoittaja ja Stakes

Suomen kielen tarkistus: Terhi Heikkonen

Ruotsinkielinen tiivistelmä: Marina Andersson

Englanninkielinen tiivistelmä: Leena Saarela

Kannen suunnittelu: Harri Heikkilä

Taitto: Christine Strid

ISBN 951-33-1777-3 (painettu julkaisu)

ISSN 1236-0732

ISBN 951-33-1661-0 (verkkokirja)

Gummerus Kirjapaino Oy

Saarijärvi 2005

Kiitokset

Tärkeä vaihe työelämässä on usean vuoden työn jälkeen päättynyt. Perehtyminen itselleni uuteen tieteenalaan on ollut haastavaa ja jälkikäteen voi todeta väitöskirjan tekemisen olleen monelta osin hyvin antoisaa. Kiinnostukseni tutkimuksen tekemiseen heräsi sosiaalipolitiikan opintojeni viimeisinä vuosina Turun yliopistossa. Edesmenneen professori Kari Salavuon innostus sosiaalipalvelujen tutkimukseen ja sosiaalipolitiikan taloudellisiin kysymyksiin antoi hyvän pohjan työlleni Stake-sin terveystaloustieteen keskuksessa.

Ohjaajaani professori Hannu Valtosta haluan kiittää perehdyttämisestä taloustieteelliseen ajattelutapaan ja ylipäättään tutkimuksen tekemiseen tutkijaurani alkutaipaleella. Hannun laaja-alainen suhtautuminen sosiaali- ja terveystaloudellisten kysymysten tutkimiseen on ollut minulle hyvänä esimerkkinä. Kiitän myös kannustuksista ja rohkaisuista erityisesti urani alkuvaiheessa sekä varauksettomasta tuesta työurani kriittisissä vaiheissa.

Toista ohjaajaani tutkimusprofessori Unto Häkkistä kiitän mahdollisuudesta väitöskirjan tekemiseen. Unton ympyrätalon käytävällä aikanaan tarjoamaan mahdollisuuteen halusin tarttua. Unton kanssa olen myös vuosien varrella käynyt lukuisia keskusteluja tilastollisista menetelmistä ja niihin liittyvistä yksityiskohdista. Nämä keskustelut ovat edesauttaneet merkittävästi ammatillista kehittymistäni.

Työni esitarkastajat professori Reijo Tilvis ja professori Harri Sintonen antoivat käsikirjoitukseen hyödyllisiä kommentteja ja parannusehdotuksia, joista olen kiitollinen. Kiitän myös mielenkiintoisista keskusteluista.

Vanhusten laitoshoidon taloustieteellinen tutkimus ei olisi mahdollista ilman ymmärrystä itse tutkimuskohteesta. Vanhustenhoidon moninaiseen maailmaan minut ovat perehdyttäneet erityisesti THT Anja Noro ja LT Harriet Finne-Soveri. Aiempi koulutukseni lasten päivähoitajaksi tarjosi välineitä ja ymmärrystä elämänsä alkutaipaleella olevien lasten kanssa työskentelyyn. Anjan ja Harrietin avulla työskennellessäni RAI-hankkeessa olen oppinut tuntemaan, mitä tämän taipaleen toisessa päässä on. Anjaa kiitän myös luottamuksesta ja jo tutkijauran alkuvaiheessa annetuista vastuullisista tehtävistä.

FT Magnus Björkgren on ollut tärkeä apuna ja hänen tutkimuksensa hyvänä esimerkkinä omalle tutkimukselleni. Esh Kaija Lindmania ja TtM Tarja Itkosta kiitän myös perehdyttämisestä vanhustenhoidon kysymyksiin ja avusta aineistojen keräämisen suunnittelussa. HSO Marianna Savolainen on avustanut tietojen keräämisessä ja huolehtinut hyvin sujuneista käytännön järjestelyistä. VTK Satu Kerpilä on tehnyt tärkeää työtä tutkimuksessa tarvittavien aineistojen kokoamisessa.

Professori Gunnar Rosenqvist ja VTM Reijo Sund ovat olleet aina käytettävissä kun olen kiireen keskellä tarvinnut tilastotieteellistä konsultaatiota. Christine

Stridiä kiitän työn taittamisesta. Tutkimukseen on saatu erillisrahoitusta Suomen Akatemiasta (päättö 70034), Sosiaali- ja terveyshallinnon ja -talouden valtakunnallisesta tutkijakoulusta ja Stakesista. Ilman rahoittajien tukea ja uskoa tutkimukseen työ ei olisi ollut toteutettavissa.

Erityisen kiitokseni haluan esittää dosentti Miika Linnalle. Hän on ollut tuttu vieras ovellani. Miikan lähes jokapäiväiset kysymykset ”No, miten sujuu, onko artikkeli edennyt, mitä tehdään seuraavaksi?” ovat pitäneet minut sopivan kiireisenä. Kiitän Miikaa myös erinomaisesta perehdyttämisestä taloustieteelliseen tuottavuustutkimukseen ja menetelmällisestä avusta. Työskentely kanssasi on ollut moneltakin osin tehokasta.

Lopuksi haluan kiittää vaimoani Terhiä. Terhin tuki ja kannustus on ollut tärkeää. Ilman sitä en ehkä olisi aikanaan päätenyt edes yliopistoon opiskelemaan. Terhi on tehnyt korvaamatonta työtä myös tarkistamalla käsikirjoituksen kieliasun. Työelämä on tärkeää, mutta elämässä kaikkein tärkeimmät asiat ovat kuitenkin muualla. Terhi ja hieman yli yksivuotias esikoisemme Kielo ovat pitäneet tutkijan maanpinnalla ja kiinni arkielämässä. Onneksi.

Helsingissä huhtikuussa 2005

Juha Laine

Tiivistelmä

Juha Laine. Laatu ja tuotannollista tehokkuutta? Taloustieteellinen tutkimus vanhusten laitoshoidosta. Stakes, Tutkimuksia 151. Helsinki 2005.

ISBN 951-33-1777-3

Tutkimuksessa tarkastellaan vanhusten pitkäaikaisen laitoshoidon tuotannollisen tehokkuuden ja hoidon laadun välistä yhteyttä. Tuotannollinen tehokkuus on tässä tutkimuksessa tuottavuuteen eli tuotos/panos-suhteeseen liittyvä käsite. Se tarkoittaa yksikön tuottavuutta suhteessa parhaimpaan mahdolliseen tuottavuuteen. Tutkimuksessa tarkastellaan aluksi vanhusten pitkäaikaisen laitoshoidon erityisluonnetta tuotannollisen tehokkuuden ja laadun käsitteellistämisen ja mittaamisen näkökulmasta. Tämän jälkeen arvioidaan, kuinka hyvin vanhainkotien ja terveyskeskussairaaloiden osastot onnistuvat optimoimaan voimavaransa, joita käytetään hoitopäivien tuotantoa varten. Lisäksi arvioidaan, miten tuotannollinen tehokkuus ja hoidon laatu ovat yhteydessä toisiinsa ja kuinka merkittävä vaihtosuhte (trade-off) niiden välillä on. Lopuksi tarkastellaan, mitkä osasto- ja yksilötason tekijät selittävät pitkäaikaishoidossa olevan asiakkaan saamaa hoitoaikaa ja millaisia eroja yksiköiden välillä on henkilöstövoimavarojen käytössä.

Tutkimus on toteutettu Sosiaali- ja terveysalan tutkimus- ja kehittämiskeskuksessa (Stakes) osana RAI-tietojärjestelmän käyttöönotto ja pitkäaikaishoidon benchmarking-hanketta. Tutkimuksessa hyödynnetään hankkeessa kertyneitä aineistoja, joiden avulla perinteiseen tuotannollisen tehokkuuden mittaamiseen voidaan liittää myös tuotoksen ominaisuuksia ja toiminnan laatua kuvaavia tekijöitä. Tutkimus perustuu yksilö- ja osastotason tietoihin ja pääosin vuosien 2001 ja 2002 aineistoihin.

Vanhusten pitkäaikaisen laitoshoidon tuottavuus on laskenut vuodesta 2000 vuoteen 2003 noin seitsemän prosenttia. Toisin sanoen palveluihin käytettyjen voimavarojen suhde tuotettujen hoitopäivien määrään on heikentynyt. Osastojen keskimääräinen tuotannollinen tehottomuus oli noin 20 prosenttia. Tämä tarkoittaa, että osastot käyttivät keskimäärin viidenneksen enemmän henkilöstö- ja pääoma-voimavaroja asiakasrakenteen avulla vakioitujen hoitopäivien tuottamiseen kuin tuotannollisesti kaikkein tehokkaimmat yksiköt. Myös voimavaroihin, prosesseihin ja tuloksiin liittyvässä hoidon laadussa oli merkittäviä eroja osastojen välillä. Lisäksi henkilöstön sijoittelu osastoille oli osittain epätarkoituksenmukainen.

Tuotannollisen tehokkuuden ja hoidon laadun välillä osoitettiin olevan vaihtosuhte. Osaston korkeaan tuotannolliseen tehokkuuteen olivat pääasiassa yhteydessä sellaiset laatuindikaattorit, jotka kuvaavat passivoivaa ja ei-kuntouttavaa hoitotyötä tai tällaisen hoitotyön ja -käytännön tuloksia. Tämän lisäksi alhainen henkilöstö- ja pääomaresursointi olivat yhteydessä korkeaan tuotannolliseen te-

hokkuuteen. Tutkimustulosten mukaan laadun yhteys tuotannolliseen tehokkuuteen ei kuitenkaan ole erityisen voimakas. Pienet muutokset hoidon laadussa eivät välttämättä vaikuta merkittävästi tuotannolliseen tehokkuuteen, eivätkä pienet muutokset tuotannollisessa tehokkuudessa vaikuta merkittävästi laatuun. Laadun ja tuotannollisen tehokkuuden välinen yhteys on merkittävämpi vasta silloin, kun näissä tekijöissä tapahtuu suuria muutoksia. Sellaisiin yksiköihin tulee kiinnittää erityistä huomiota, jotka ovat joko hoidon laadun tai tuotannollisen tehokkuuden suhteen muista selvästi poikkeavia.

Laitoshoidon tuottavien yksiköiden toimintaa tulisi tutkimuksen mukaan arvioida aina useasta eri näkökulmasta ja usean eri mittarin avulla. Laitoshoidossa on kehittämistarpeita sekä tuotannollisen tehokkuuden, työvoiman käytön että hoidon laadun suhteen, joita ei voida ratkaista vain lisäämällä henkilöstön määrää tai parantamalla tuottavuutta. Olennaista onkin, millaiseksi tuotannon ja voimavarojen määrän ja laadun välinen suhde päätöksenteossa muodostuu. Tätä ei voida tarkastella empiirisesti ilman käyttökelpoisia seurantajärjestelmiä ja mittareita. Vaikka vanhusten laitoshoidon tuottavuuden parantamiseen ei toiminnan erityisen luonteen takia ole tulevaisuudessa merkittäviä mahdollisuuksia, voivat yksittäiset yksiköt parantaa omaa tuotannollista tehokkuuttaan voimavarojen aiempaa onnistuneemmalla käytöllä. Tällöin on kuitenkin varmistettava, ettei hoidon laatu kehity merkittävästi huonompaan suuntaan.

Tutkimus tarjoaa lähestymistapoja sekä palvelutuottajien toiminnan että erityyppisten hoivapalvelujen arviointiin. Tutkimustuloksia voidaan hyödyntää myös arvioitaessa erityyppisten hoivapalvelujen tuottavuuden parantamiseen liittyviä riskejä.

Avainsanat: tuotannollinen tehokkuus, tuottavuus, laatu, vanhusten laitoshoido, RUG-III, RAI

Sammandrag

Juha Laine. Laatu ja tuotannollista tehokkuutta? Taloustieteellinen tutkimus vanhusten laitoshoidosta [Kvalitet och produktionseffektivitet? En ekonomisk studie om institutionsvård för äldre]. Stakes, Tutkimuksia 151. Helsingfors 2005. ISBN 951-33-1777-3

I studien analyseras kopplingen mellan produktionseffektivitet och vårdkvalitet inom långvarig institutionsvård för äldre. Produktionseffektivitet är i denna studie ett begrepp som hänger samman med produktiviteten, dvs. förhållandet mellan insatsen och utbytet. Det betyder enhetens produktivitet i förhållande till den bästa möjliga produktiviteten. I studien analyseras först den särskilda karaktären hos långvarig institutionsvård för äldre när man utgår från bildning av begrepp för och mätning av produktionseffektivitet och kvalitet. Efter detta bedöms hur väl ålderdomshemmens och hälsocentralsjukhusens avdelningar lyckas optimera sina befintliga resurser för produktion av vård dagar. Dessutom bedömer man hur produktionseffektiviteten och vårdkvaliteten är kopplade till varandra och hur signifikant bytesrelationen (trade-off) mellan dem är. Slutligen analyseras vilka faktorer på avdelnings- och individnivå som förklarar den vårdtid som ägnats långvårdsklienten.

Studien genomfördes vid Forsknings- och utvecklingscentralen för social- och hälsovården (Stakes) som en del av projektet att ta i bruk informationssystemet RAI (Resident Assessment Instrument) och utveckla ett system för jämförelse (benchmarking) inom långvården. I studien utnyttjades data som erhållits i projektet. Genom dem kan man också bifoga faktorer som avspeglar resultatets egenskaper och verksamhetskvaliteten till den traditionella mätningen av produktionseffektivitet. Studien bygger på data på individ- och avdelningsnivå från främst 2001 och 2002.

Produktiviteten inom långvarig institutionsvård för äldre har minskat cirka sju procent från 2000 till 2003. Med andra ord har förhållandet mellan de resurser som satsats på service och antalet producerade vård dagar försvagats. Avdelningarnas genomsnittliga produktionsineffektivitet var cirka 20 procent. Det betyder att avdelningarna använde i snitt en femtedel mer personal- och kapitalresurser för att producera vård dagar standardiserade enligt klientstrukturen än de produktionsmässigt mest effektiva enheterna. Även vårdkvaliteten när det gäller kraftresurser, processer och resultat varierade avsevärt mellan olika avdelningar. Dessutom var fördelningen av personal på avdelningarna delvis oändamålsenlig.

Man kunde påvisa att det fanns en bytesrelation mellan produktionseffektiviteten och vårdkvaliteten. Särskilt sådana kvalitetsindikatorer som beskriver passiverande och icke rehabiliterande vårdarbete eller resultat av vårdarbete och vårdpra-

xis av denna typ hängde samman med hög produktionseffektivitet på avdelningen. Utöver detta hängde även tilldelning av små personal- och kapitalresurser samman med hög produktionseffektivitet. Enligt studieresultaten var emellertid kopplingen mellan låg vårdkvalitet och hög produktionseffektivitet inte särskilt stark. Små ändringar i vårdkvaliteten påverkar inte nödvändigtvis produktionseffektiviteten i någon större utsträckning, liksom inte heller små förändringar i produktionseffektiviteten påverkar kvaliteten i någon större utsträckning. Kopplingen mellan kvalitet och produktionseffektivitet är av större betydelse först när dessa förändras mycket. Enheter som avviker från andra när det gäller vårdkvalitet eller produktionseffektivitet bör dock särskilt uppmärksammas.

Verksamheten i enheter som producerar institutionsvård bör enligt studien bedömas ur flera olika perspektiv och genom flera olika instrument. Inom institutionsvården finns utvecklingspotential både när det gäller produktionseffektivitet, användning av arbetskraft och vårdkvalitet. Väsentligt är hurdan kopplingen mellan resurser, produktionsvolymen och kvaliteten inom beslutsfattandet är. Detta kan inte undersökas empiriskt utan ändamålsenliga uppföljningssystem och instrument. Även om det på grund av verksamhetens särskilda karaktär inte finns särskilt stora möjligheter till att förbättra produktiviteten inom institutionsvård för äldre, kan enskilda enheter förbättra sin produktionseffektivitet genom att använda sina resurser på ett mer lyckat sätt än tidigare. I samband med detta bör man dock säkerställa att vårdkvaliteten inte signifikant försämras.

Genom studien fås olika tillvägagångssätt för bedömning av både serviceproducenternas verksamhet och olika typer av omvårdnadstjänster. Man kan också dra nytta av studieresultaten när olika risker och möjligheter i anslutning till förbättring av omvårdnadstjänsternas produktivitet bedöms.

Nyckelord: produktionseffektivitet, produktivitet, kvalitet, institutionsvård för äldre, RUG-III, RAI

Abstract

Juha Laine. Laatus ja tuotannollista tehokkuutta? Taloustieteellinen tutkimus vanhusten laitoshoidosta [Quality and Productive Efficiency? An Economic Study on Institutional Care for Older People]. National Research and Development Centre for Welfare and Health (STAKES), Research Report 151. Helsinki 2005. ISBN 951-33-1777-3

The study focuses on the connection between productive efficiency and care quality in institutional care for older people. The concept of productive efficiency deals with productivity, in other words, the relationship between inputs and outputs, with unit's productivity considered in relation to the best possible productivity. Firstly, we investigate the special character of institutional care for older people in regard to conceptualising and measuring productive efficiency and quality of care. After this, we evaluate the success of wards of residential homes for older people and of public health centres in optimising their resources for producing inpatient days. Additionally, we evaluate the connection between productive efficiency and the quality of care and ask whether the trade-off between them is significant. Finally, we analyse which ward- and individual-level factors explain the amount of care time given to a long-term care client.

The study was made at the National Research and Development Centre for Welfare and Health (STAKES) as part of the project 'Benchmarking and the Implementation of the RAI System in Elderly Care'. The data from this project made it possible for us to add output and quality characteristics to the traditional measures of productive efficiency. Individual- and ward-level data mainly from 2001 and 2002 were used.

The productivity of institutional care for older people in 2000–2003 declined 7 per cent. In other words, the ratio between the resources used and the number of produced inpatient days has diminished. The average productive inefficiency of the ward was approximately 20 per cent. This implies that the wards used on average a fifth more personnel and capital resources to produce the same amount of case-mix adjusted inpatient days as the most efficient wards. Even the quality of care in terms of resources, processes and outcomes differed notably between the wards. Furthermore, the placement of personnel on the wards was partly inappropriate.

The study demonstrated a trade-off between productive efficiency and the quality of care. The high productive efficiency at ward level was mainly associated with quality indicators that reflect passivating and non-rehabilitating care or the outcomes of such care practices. Moreover, a low level of personnel and capital resourcing was associated with a high productive efficiency. However, the results show the connection between quality and productive efficiency was not especially

strong. Small changes in the quality of care do not necessarily have a major influence on productive efficiency or vice versa. The connection between quality and productive efficiency becomes of greater importance when these change remarkably. Units that differ from others in terms of quality of care or productive efficiency should be given special attention.

Further, the activities of units producing institutional care should be evaluated from different viewpoints and by different indicators. In institutional care, there is development potential for productive efficiency, the use of labour and the quality of care. The essential issue in the decision-making is how the relationship between the resources, the amount and the quality of production is determined. This cannot be studied empirically without appropriate monitoring systems and measures. Even though there are no significant possibilities to improve the productivity of institutional care for older people because of the special character of the activity, individual units can improve their productive efficiency by a more successful use of resources. However, in this case it should be made sure that the care quality does not significantly deteriorate.

The study presents ways of approaching the evaluation of both service producer activity and care services. The study results can also be of use when evaluating risks connected to improving productivity of different kinds of care services.

Keywords: productive efficiency, productivity, quality, institutional care for older people, RUG-III, RAI

Sisällys

Kiitokset

Tiivistelmä

Sammandrag

Abstract

Taulukko- ja kuvioluettelo	13
Lyhenteet.....	14
Alkuperäisjulkaisut	15
1 Johdanto	17
2 Vanhustenhuollon muuttuva toimintaympäristö	21
3 Vanhusten laitoshoidon erityispiirteet taloustieteen näkökulmasta	24
3.1 Toimialan erityinen luonne	24
3.2 Tiedon epäsymmetria.....	24
3.3 Toiminnan laadullisten olottuvuuksien arvioiminen	25
3.4 Hintajärjestelmän poikkeavuus	26
3.5 Asiakkaan rajoitettu valinnanmahdollisuus	27
3.6 Budjettirajoitteet ja tulonmuodostus	27
3.7 Ihmistyön korvaamattomuus.....	28
4 Palvelutuotantoprosessi, tuotannollinen tehokkuus ja laatu	29
4.1 Terveyspalvelujen tuotantoprosessi ja taloustieteellinen tuottavuustutkimus..	29
4.2 Laadun käsite.....	31
4.3 Laadun olottuvuudet.....	33
4.4 Laadun olottuvuuksien mittaaminen	35
4.5 Palvelutuottajien tavoitteet ja kannustimet	37
4.6 Tuotannon määrän ja laadun vaihtosuhde	39
5 Tuottavuuden ja tuotannollisen tehokkuuden mittaaminen	41
5.1 Terveystenhuollon tuottavuuden mittaaminen.....	41
5.2 Parametrinen lähestymistapa.....	42
5.3 Epäparametrinen lähestymistapa.....	44
6 Tuottavuuserojen tutkimus.....	46
6.1 Tuotannollisen tehokkuuden erojen selittäminen	46
6.2 Vanhustenhuollon tuotannollisen tehokkuuden ja laadun empiiriset tutkimukset	48
7 Tutkimuksen tavoitteet.....	52
8 Tutkimusaineistot.....	54
8.1 Tietolähteet	54
8.2 Havaintoyksiköt	55
8.3 Asiakasrakenne ja voimavaravaativuus	57

- 8.4 Tuotoksen mittaaminen 58
- 8.5 Kustannustiedot ja panoshinnat 60
- 8.6 Tuotannon tekijät 60
- 8.7 Aikamittaus 61
- 8.8 Muut tiedot 62
- 8.9 Laatumuuttajat 63
- 9 Tutkimusmenetelmät 65
 - 9.1 Parametriset menetelmät 65
 - 9.1.1 Stokastiset rintamamallit ja monitasomallinnus 65
 - 9.1.2 Malliin liittyvien oletusten ja mallispesifikaation testaaminen 66
 - 9.2 Epäparametriset menetelmät 68
 - 9.2.1 Data envelopment -analyysi 68
 - 9.2.2 Hyvän ja heikon hoidon laadun osastojen vertailu tuotannollisen tehokkuuden suhteen 69
- 10 Tulokset 70
 - 10.1 Pitkäaikaishoidon tuotannollinen tehokkuus (osajulkaisut 1, 2 ja 3) 70
 - 10.2 Hoidon laadun yhteys tekniseen tehokkuuteen (osajulkaisut 1 ja 2) 71
 - 10.3 Hoidon laadun yhteys kustannustehokkuuteen ja kustannuksiin (osajulkaisu 3) 73
 - 10.4 Muita tuotannolliseen tehokkuuteen yhteydessä olevia tekijöitä (osajulkaisut 2 ja 3) 74
 - 10.5 Työajan käyttö osastoissa (osajulkaisu 4) 75
 - 10.6 Työajan kohdentumista selittävät tekijät (osajulkaisu 4) 76
- 11 Pohdinta 79
 - 11.1 Laitoshoidon tuotannollinen tehokkuus ja hoidon laatu 79
 - 11.2 Tuotannollisen tehokkuuden erojen kaventaminen ja tuottavuuden parantaminen 80
 - 11.3 Henkilöstön sijoittelun ja työajan käytön kehittäminen 81
 - 11.4 Tuotannollinen tehokkuus ja etiikka 83
 - 11.5 Tutkimukseen liittyvät varaukset ja jatkotutkimustarpeet 84
- Lähteet 86
- Liite 1. Tutkimuksessa käytetyt kliiniset laatuindikaattorit ja niiden laskenta MDS 2.0-lomakkeesta 91

Taulukko- ja kuvioluetelo

Taulukot

Taulukko 1.	Laadun arvioinnin malli laadun ulottuvuuksien ja arvioijan mukaan	34
Taulukko 2.	Osajulkaisuissa käytetyt aineistot, muuttajat ja menetelmät	56
Taulukko 3.	Teknisen tehokkuuden ero hyvän ja huonon hoidon laadun osastoissa	72
Taulukko 4.	Yksittäiselle asiakkaalle kohdentunut työaika ja työajan muutos vuodesta 1995 vuoteen 2002	76
Taulukko 5.	Hoitohenkilökunnan palkoilla painotettu hoitoaika minuuteissa vuorokauden aikana asiakkaiden fyysisen toimintakyvyn (ADL) ja kognition (CPS) mukaan	77

Kuviot

Kuvio 1.	Vanhusväestön väestöennuste vuosille 2000–2030	21
Kuvio 2.	Pitkäaikaissosastojen tuottavuuden kehitys vuosina 2000–2003	23
Kuvio 3.	Terveystenhoitotoiminnan elementit ja tehokkuustutkimuksen osatekijät ...	29
Kuvio 4.	Asiakkaan fyysisen, psyykkisen, kognitiivisen tai sosiaalisen toimintakyvyn kehitys laitoshoidossa	33
Kuvio 5.	Laadun ja määrän vaihtosuhdekäyrä	39
Kuvio 6.	Tekninen tehokkuus ja tehokkuusrintama	45
Kuvio 7.	RUG-III/22-luokitusjärjestelmä ja suomalaiset kustannuspainot	58
Kuvio 8.	Taloussyhdyskenkilöiltä ja RAI-tietokannasta saatujen hoitopäivien välinen yhteys	59
Kuvio 9.	Hoidon laatu jatkuvana ja luokittelumuuttujana	64
Kuvio 10.	Tehokkuuslukujen jakaumat osajulkaisuissa 1–3	70
Kuvio 11.	Teknisen tehokkuuden ja osaston hoitopaikkojen määrän välinen yhteys	74
Kuvio 12.	Mittakaavaedun (skaalatehokkuus) ja osaston hoitopaikkojen määrän välinen yhteys	75
Kuvio 13.	Perinteinen ja asiakasrakenteella painotettu henkilöstömitoitus	82

Lyhenteet

15D	Suomessa kehitetty terveyteen liittyvä geneerinen elämänlaadun mittari
ADL	Fyysinen toimintakyky, päivittäiset toiminnot (Activities of daily living)
COLS	Korjattu pienimmän neliösumman regressioanalyysi (Corrected ordinary least square regression)
CPS	Kognitiivisten toimintojen mittari (Cognitive performance scale)
CRS	Vakioskaalatuotot (Constant returns to scale)
DEA	Data Envelopment Analysis on epäparametrinen menetelmä, jota käytetään tehokkuuslukujen laskemiseen
MDS 2.0	Minimum Data Set -lomakkeiston versio 2.0. Vähimmäistietomäärä, joka tarvitaan pitkäaikaishoidossa olevan asiakkaan arviointia ja hoidon tarpeen kartoitusta varten.
PNS	Pienimmän neliösumman regressioanalyysi (Ordinary least squares regression)
QALYs	Laatupainotteiset elinvuodet (Quality adjusted life years)
RAI	Yhdysvalloissa kehitetty asiakkaan arviointijärjestelmä (Resident Assessment Instrument)
RUG-III	Asiakasrakenneluokitus, jota käytetään asiakkaiden voimavara-tarpeen arvioitiin (Resource utilization groups)
RUG-III/22	RUG-III-luokitusjärjestelmän suomalainen 22-luokkainen versio
SAS-toiminta	Selvitä, Arvioi, Sijoita -toiminnan tavoitteena on arvioida eri hoidontasoille jonottavien asiakkaiden elämäntilannetta ja sijoittaa heidät sopivalle hoitopaikalle
SCE	Skaalatehokkuus (Scale efficiency)
VRs	Vaihtuvat skaalatuotot (Variable returns to scale)

Alkuperäisjulkaisut

- I Laine, J., Finne-Soveri, H., Björkgren, M., Linna, M., Noro, A. & Häkkinen, U. (2005). The Association between Quality of Care and Technical Efficiency in Long-Term Care. *International Journal for Quality in Health Care* 17(3): 259–267.
- II Laine, J., Linna, M., Häkkinen, U. & Noro, A. (2005). Measuring the Productive Efficiency and Clinical Quality of Institutional Long-Term Care for the Elderly. *Health Economics* 14(3): 245–256.
- III Laine, J., Linna, M., Noro, A. & Häkkinen, U. 2005. The Cost Efficiency and Clinical Quality of Institutional Long-term Care for the Elderly. *Health Care Management Science* 8(2): 149–156.
- IV Laine J., Noro, A., Finne-Soveri, H. & Häkkinen, U. (2005). Patient- and ward-level determinants of nursing time in nursing facilities. *Journal of Health Services Research & Policy* (in press).

1 Johdanto

Sosiaali- ja terveysalalla kiinnostus palvelujen taloudelliseen tarkasteluun oli pitkään vähäistä. Hyvinvointivaltion rakentamisen aikana painopiste oli pikemminkin palvelu- ja sosiaaliturvajärjestelmän laajentamisessa, jossa valtiolla ja kunnilla oli keskeinen asema sosiaali- ja terveyspalvelujen tuottajina. Laman aikana ja sen jälkeen 1990-luvulla julkiseen keskusteluun tulivat sellaiset käsitteet kuin tilaaja–tuottaja-malli ja yksityinen palvelutuotanto, joissa painottuivat voimakkaasti palvelutuotannon talouteen liittyvät näkökulmat ja palvelujen vaihtoehtoiset tuotantotavat. Tämä keskustelu on jatkunut voimakkaana 2000-luvulle tultaessa. Viime vuosina julkisilta palveluilta ja julkisilta organisaatioilta on vaadittu aiempaa voimakkaammin tuottavuutta ja tehokkuutta. Toisin sanoen palveluihin ja tulonsiirtoihin käytettyjä voimavaroja ja niistä aiheutuvia kustannuksia tulisi arvioida suhteessa tuotokseen eli tuotettujen palvelujen määrään, jolloin puhutaan tuottavuudesta (= tuotos / panos). Toisaalta käytettyjä voimavaroja pitää arvioida myös suhteessa palvelujen tuottamiin hyvinvointi- ja terveyshyötyihin, jolloin puhutaan kustannus-vaikuttavuudesta eli tehokkuudesta.

Joillekin suomalaisille talouteen liittyvät kysymykset, kuten esimerkiksi palvelujen tuottavuus liittyy mielikuviin 1990-luvun laman aikana palveluihin ja sosiaaliturvaan tehdystä leikkauksista (ks. Heikkilä & Uusitalo, 1997). Tällöin keskustelu palvelujen tuottavuudesta voidaan kokea sopimattomaksi ja epäeettiseksi. Palvelujen tuotantoon ja järjestämiseen liittyvät taloudelliset kysymykset koetaan Suomessa nykyään kuitenkin aiempaa tärkeämmäksi. Tästä viimeisimpänä esimerkkinä on vuoden 2004 kunnallisvaalien yhteydessä virinnyt keskustelu palvelujen erilaisista tuottamistavoista ja niiden eduista ja haitoista. Syitä taloudellisen keskustelun lisääntymiseen voidaan löytää sekä sosiaali- ja terveydenhuollon sisältä että yleisistä yhteiskunnallisista ilmiöistä. Ensinnäkin sosiaali- ja terveydenhuollossa on koettu ongelmaksi niin kutsuttu Baumolin tauti (Baumol, 1967). Se tarkoittaa tilannetta, jossa palvelualan reaalipalkat ovat nousseet muita toimialoja vastaavasti, vaikka palvelualan tuottavuuskehitys on ollut heikompaa tuotannon erityisen luonteen takia.¹ Tämän seurauksena palvelujen suhteelliset hinnat eli kustannukset nousevat. Baumolin tautia voidaan yrittää hillitä parantamalla palvelualan tuottavuutta. Toiseksi väestön nopea ikääntyminen on pakottanut valtion ja kunnat pohtimaan sosiaaliturvan ja palvelujen rahoitusta ja järjestämistä. Suomessa esimerkiksi eläkejärjestelmää on muutettu merkittävästi. Kolmanneksi taloudellisen keskustelun voimistuminen liittyy hyvinvointivaltion institutionaaliin ja ideologisiin muu-

¹ Parkkisen (Parkkinen, 2004a) mukaan hoivatyössä työn tuottavuus kasvoi vuodesta 1975 vuoteen 2002 keskimäärin vain 0,5 prosenttia vuodessa, kun koko kansantaloudessa tuottavuuden parantuminen on ollut huomattavasti nopeampaa.

toksiin, joita esimerkiksi Alasuutari on kuvannut siirtymisenä suunnitelmataloudesta kilpailutalouteen (Alasuutari, 2004).

Tällä hetkellä julkisella sektorilla yksi merkittävimmistä sosiaali- ja terveyspalvelujen tuotannon talouteen liittyvistä kysymyksistä on julkisten palvelujen ja hallinnon tuottavuus. Vuonna 2003 valtionvarainministeriö käynnisti julkisen sektorin tuottavuuden toimenpideohjelman (tuottavuushankkeen), jossa tavoitteena on ”... julkisen hallinnon ja julkisten palvelujen tuottavuuden ja tehokkuuden kasvu, joka saadaan aikaan toiminnallisin ja rakenteellisin muutoksin ja tieto- ja viestintätekniiikan käytöllä, tehostamalla yleishallinnon toimintoja ja käyttämällä yksityisiä palveluja tehokkaammin julkisten palvelujen tuottamisessa” (VM, 2003). Tässä yhteydessä tehokkuudella ei kuitenkaan tarkoiteta käsittääkseni kustannus-vaikutavuutta vaan kutakuinkin samaa kuin tuottavuudella. Pääministeri Matti Vanhasen hallituksen ohjelmaan liittyen on puolestaan laadittu peruspalveluohjelma vuosiksi 2005–2008. Tähän ohjelmaan liittyy vuosittainen katsaus peruspalvelujen tuottavuuteen. Aiemmin tuottavuuden arviointi on Suomessa ollut vain osa julkisten sektoritutkimuslaitosten omaa tutkimustoimintaa, jolloin tutkimustuloksia ei välttämättä ole välittömästi hyödynnetty poliittisessa päätöksenteossa. Valtion taloudellinen tutkimuskeskus (VATT), Suomen Kuntaliitto, Sosiaali- ja terveysalan tutkimus- ja kehittämiskeskus (Stakes) sekä Tilastokeskus julkaisivat vuonna 2003 ensimmäisen laajan kunnallisten palvelujen tuottavuutta käsittelevän kirjan, jossa tarkasteltiin terveyden- ja vanhustenhuollon tuottavuutta Suomessa (Hjerpe, Kangasharju & Vuorento, 2003). Palvelujen tuottavuus on viime vuosina ollut terveystaloustieteellisen tutkimuksen kohteena aiempaan useammin, joten sekä puhtaasti tieteellisten että poliittista päätöksentekoa tukevien tuottavuustutkimusten määrän voidaan odottaa lisääntyvän lähivuosina.

Tuottavuuden painoarvo on ymmärrettävää, sillä onhan tuottavuus kansantalouden tärkeä menestymisen mittari. Tuottavuudella kuvataan kansankunnan kilpailukykyä ja se nähdään myös keinona saavuttaa korkea elintaso. Tuottavuuden kasvu turvaa myös palvelujen ja sosiaaliturvan rahoitusta. Sosiaali- ja terveyspalvelujärjestelmän tavoitteet ovat kuitenkin huomattavasti tuottavuutta moninaisemmat, jolloin sen rinnalle tarvitaan muita arviointikriteereitä ja tavoitteita. Palvelujärjestelmän kuten tulonsiirtojärjestelmänkin yhteiskunnalliset tavoitteet liittyvät yleisesti tehokkuuden ja oikeudenmukaisuuden tavoittelemiseen (Barr, 1998; Le Grand, Proper, & Robinson, 1992). Tehokkuudella tarkoitetaan tällöin voimavarojen optimaalista käyttöä hyvinvoinnin, terveyden ja elämänlaadun aikaansaamiseksi. Oikeudenmukaisuudella tarkoitetaan vuorostaan palvelujen jakautumista kansalaisille joidenkin yhteisesti hyväksytyjen oikeudenmukaisuusperiaatteiden mukaan. Palvelutuotannon eri organisointitavoilla pyritään saavuttamaan nämä tavoitteet. Palvelutuotannossa joudutaan siten vastaamaan kysymyksiin siitä mitä, miten ja kenelle palvelut tuotetaan. Tehokkuuden ja oikeudenmukaisuuden tavoitteet saattavat kuitenkin olla ristiriidassa (Okun, 1975) eikä niiden yhtäaikainen

saavuttaminen ole helppoa tai välttämättä edes mahdollista. Sosiaali- ja terveysalalla tuottavuutta korostavan taloudellisen puheen koettu uhka näyttäytyy huolena sekä oikeudenmukaisuuden toteutumisesta että erityisesti pelkona palvelujen laadun ja inhimillisen kohtelun katoamisesta.

Väestön ikääntyminen, vanhusten hoivapalveluihin liittyvä tiedon epäsymmetria ja laadun arvioimisen vaikeus merkitsevät, että vanhusten hoivapalveluja tuottavien organisaatioiden tuottavuuden ja laadun arviointi on tärkeää. Erityisesti nykyisessä yhteiskunnallisessa tilanteessa, jossa yhteiskuntapolitiikan julkilausuttuna tavoitteena on palvelujen tuottavuuden parantaminen. Yhteiskunnallisesta näkökulmasta tarkasteltuna tehottomuuden eli heikon kustannusvaikuttavuuden hinta terveydenhuollossa on pienempi hoidon kokonaismäärä, korkeammat kustannukset ja suuremmat hyvinvointitappiot. Korkean tuottavuuden hintana saat-
taa puolestaan olla inhimillisesti katsottuna huono hoito. Toistaiseksi on tehty kuitenkin niukasti tutkimuksia siitä, mitkä ovat tuottavuuden lisäämisen riskit ja mahdolliset seuraukset erityyppisissä hoivapalveluissa. Esimerkiksi käsitykset vanhusten laitoshoidon tuottavuuden ja hoidon laadun välisestä yhteydestä eivät ole yhdenmukaisia. Laatu on moniulotteinen ilmiö, jonka mittaaminen ja sisällyttäminen kattavasti tuottavuustutkimuksiin on välttämätöntä.

Tämä tutkimus kuuluu taloustieteen tuottavuustutkimuksen alaan, jossa tuottavuuden ja laadun välistä yhteyttä voidaan pitää eräänä peruskysymyksenä. Tässä tutkimuksessa tutkimuskohdetta eli vanhusten pitkäaikaista laitoshoidoa tarkastellaan taloustieteellisen tuottavuustutkimuksen käsittein ja menetelmin. Vanhusten laitoshoidon tuotantoprosessi käsitteellistetään kvantitatiivisena prosessina, jossa tuotannontekijöillä tuotetaan tuotoksia eli suoritteita. Tuotannollinen tehokkuus on tässä tutkimuksessa kvantitatiivinen suure, jolla mitataan palvelua tuottavan yksikön tuottavuutta suhteessa parhaimpaan mahdolliseen tuottavuuteen. Tässä tutkimuksessa tuotannollisella tehokkuudella ei siten tarkoiteta toiminnan hyvinvointi-, elämänlaatu- ja terveysvaikutuksia suhteessa voimavaroihin tai niistä aiheutuviin kustannuksiin. Tuotannollista tehokkuutta voidaan pitää tärkeänä ilman, että on mitattu toiminnan lopullisia vaikutuksia, jos toiminnan oletetaan sinänsä olevan tärkeää ja hyödyllistä. Pitkäaikaisen laitoshoidon voidaan pääasiassa olettaa johtavan siihen, että asiakkaan terveys ja hyvinvointi ovat parempia verrattuna tilanteeseen, jossa asiakas ei pitkäaikaishoidon tarpeesta huolimatta olisi palvelun piirissä. Myös palvelun laatua käsitellään tässä tutkimuksessa kvantitatiivisena suurena kun se otetaan eksplisiittisesti tarkasteluun mukaan. Tutkimuksessa tarkastellaan aluksi laitoshoidon erityisluonnetta tuotannollisen tehokkuuden ja laadun käsitteellistämisen ja mittaamisen. Tämän jälkeen arvioidaan, kuinka hyvin vanhainkotien ja terveyskeskussairaaloiden osastot onnistuvat optimoimaan käytettävissä olevat voimavaransa, joita käytetään hoitopäivien tuotantoa varten. Lisäksi arvioidaan, miten tuotannollinen tehokkuus ja hoidon laatu ovat yhteydessä toisiinsa ja kuinka merkittävä vaihtosuhde niiden välillä on. Lopuksi tarkas-

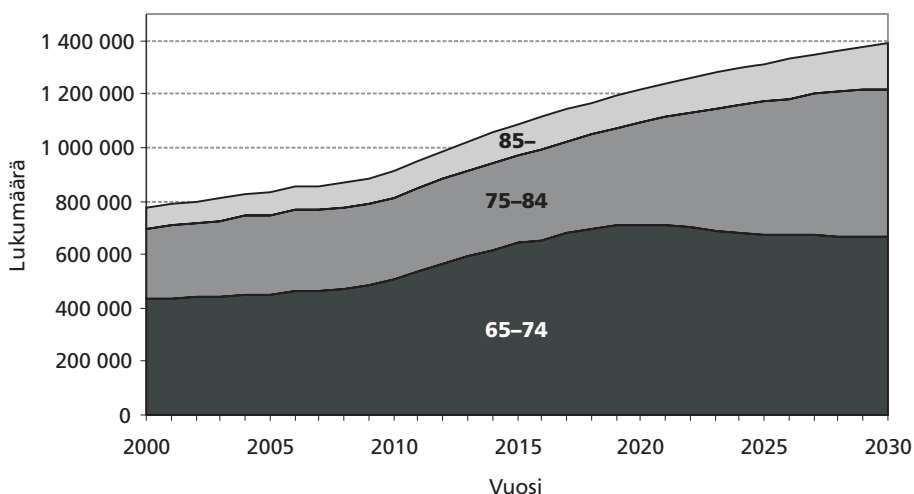
tellaan, mitkä osasto- ja yksilötason tekijät selittävät asiakkaan saamaa hoitoaikaa (min/vrk) ja millaisia eroja yksiköiden välillä on henkilöstövoimavarojen käytössä.

Tutkimuksen toisessa luvussa käsitellään olennaisimpia vanhustenhuollon toimintaympäristössä tapahtuneita ja tapahtuvia muutoksia, jotka vaikuttavat vanhuspalvelujen organisointiin ja yhteiskunnalliseen painoarvoon. Kolmannessa luvussa arvioidaan taloustieteen näkökulmasta vanhusten laitoshoidon erityispiirteitä ja niiden vaikutuksia tuotannolliseen tehokkuuteen ja hoidon laatuun. Neljännessä luvussa tarkastellaan palvelutuotantoprosessin, tuottavuustutkimuksen ja hoidon laadun käsitteitä, palvelutuottajien tavoitteita ja kannustimia sekä palvelun määrään ja laatuun liittyvää vaihtosuhdetta. Viidennessä luvussa esitellään tuottavuuden mittaamismenetelmiä. Kuudennessa luvussa tarkastellaan, millä tavoin yksiköiden tuottavuuseroja voidaan selittää ja luodaan katsaus alan kirjallisuuteen. Luvussa seitsemän esitetään tutkimuksen tavoitteet ja tutkimuskysymykset. Luvussa kahdeksan käydään läpi tutkimusaineistot ja tutkimuksessa käytetyt muuttujat. Luvussa yhdeksän esitetään tutkimusmenetelmät. Tutkimustulokset raportoidaan luvussa kymmenen ja yhteenveto ja johtopäätökset esitetään luvussa yksitoista.

2 Vanhustenhuollon muuttuva toimintaympäristö

Suomen väestön keski-ikä on kohoamassa. Tilanne on vastaava koko Euroopassa. Suomessa sekä miesten että naisten elinajanodotteen ennustetaan nousevan merkittävästi tulevina vuosikymmeninä. On arvioitu, että vuoteen 2030 mennessä yli 65-vuotiaiden osuus väestöstä kasvaa tämän hetkisestä 15 prosentista 26 prosenttiin. Määrällisesti tämä tarkoittaa noin 600 000 ihmistä (SVT, 2003). Myös demens-tiaa sairastavien määrän on ennustettu lisääntyvän tämän hetkisestä noin 75 000 noin 128 000:een vuoteen 2030 mennessä (Vaarama & Voutilainen, 2002). Kuviossa 1 on esitetty yli 65-vuotiaiden lukumäärän kehitys Suomessa vuosina 2000–2030. Vaikka ikääntyneet ovat tulevaisuudessa aiempia ikäluokkia terveempiä ja elävät pidempään, niin ikääntymisen myötä tarve saada sosiaali- ja terveystalveluja kuitenkin lisääntyy (Vaarama & Voutilainen, 2002). Tulevaisuudessa vanhusten palveluihin ja tulonsiirtoihin joudutaankin käyttämään aiempaa enemmän voimavaroja.

Vuonna 2002 noin 13 prosenttia 65 vuotta täyttäneistä oli säännöllisten sosiaali- ja terveystalvelujen piirissä (SVT, 2003). Säännöllisillä palveluilla tarkoitetaan kotihoitoa, palveluasumista ja laitoshoidoa. Suuri osa vanhuksista ei ole säännöllisten palvelujen piirissä, vaan elävät kotona esimerkiksi täysin itsenäisesti tai omaishoitajan kanssa ja käyttävät palveluja vain tilapäisesti. Vanhainkodeissa ja



Lähde: Väestötilastot, Tilastokeskus

KUVIO 1. Vanhusväestön väestöennuste vuosille 2000–2030

terveyskeskuksissa oli vuonna 2002 pitkäaikaishoidossa noin 31 000 asiakasta. Laitoshoidossa olevien asiakkaiden määrä kuitenkin vähentyi 1990-luvun aikana, jolloin vanhustenhuollon palveluissa tapahtui rakennemuutos. Tuolloin kunnissa vähennettiin laitoshoidoa ja lisättiin sekä palveluasumista että kotihoitoa. (SVT, 2003) Toinen rakennemuutokseen liittyvä ilmiö on viime vuosina ollut vanhuksille tarjotujen yksityisten asumispalvelujen ja kotiin annettavien palvelujen lisääntyminen (Kauppinen, 2001). Yksityisten palvelujen osuus pitkäaikaista vanhainkoti- ja terveyskeskushoitoa vastaavista palveluista on kuitenkin edelleen hyvin vähäinen (SVT, 2003).

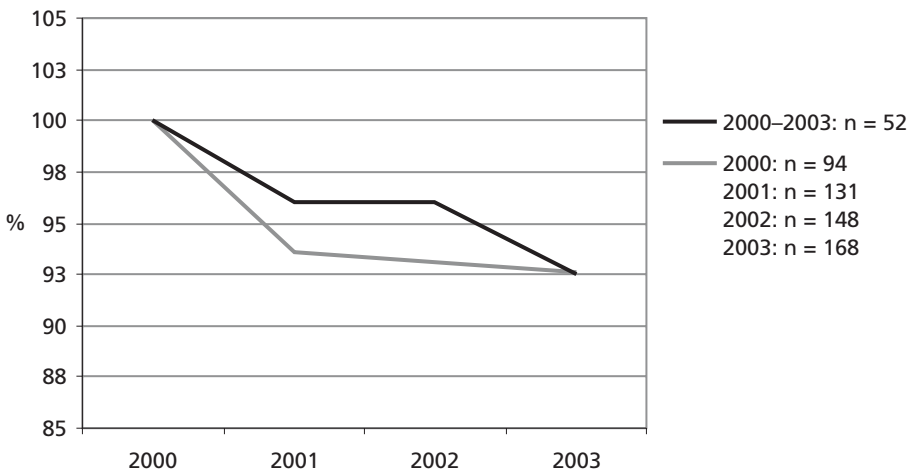
Vanhusten laitoshoido on kunnissa merkittävä työllistäjä. Viime vuosikymmenen aikana vanhainkotien henkilöstömäärä pysyi noin 19 500:ssa, mutta perusterveydenhuollon vuodeosastoilla henkilöstömäärä väheni noin 25 000:sta hieman yli 20 000:een. (SVT, 2003) Pitkäaikaishoitolaitosten henkilöstömäärä näyttäisi kuitenkin hieman lisääntyneen 2000-luvulla (Laine, 2005a). On todennäköistä, että tämä suuntaus jatkuu myös tulevina vuosina muun muassa vanhustenhuollon valtakunnallisten laatusuositusten vauhdittamana. Näiden laatusuositusten perusteella vanhusten pitkäaikaishoidossa on tarvetta lisätä henkilöstön määrää merkittävästi (Vaarama, Luomahaara, Peiponen, & Voutilainen, 2001). Toisaalta työvoiman tarjonta voi muodostua ongelmaksi. Vanhustenhuoltoa ja erityisesti työskentelyä laitoshoidossa ei välttämättä koeta houkuttelevaksi muun muassa alhaisen palkkatason ja työn raskauden johdosta. Tulevaisuudessa saatetaan tarvita nykyistä enemmän työvoimaa ulkomailta.

Vuonna 2001 vanhusten sosiaalipalvelujen menot olivat noin 1,2 miljardia euroa, mikä on noin 10 prosenttia kaikista vanhusten sosiaalipalvelujen ja tulonsiirtojen menoista. Sosiaalipalveluilla tarkoitetaan vanhainkotihoitoa, omaishoidon tukea, kotiin annettavia palveluja ja erilaisia palveluasumis- ja päivätoimintapalveluja. Vanhainkotihoitoon osuus tästä on noin 550 miljoonaa euroa. Sosiaalipalvelujen lisäksi menoihin pitää lisätä vielä vanhusten terveyskeskusvuodeosastohoito, jonka menot olivat vuonna 2001 arvioilta noin 600 miljoonaa euroa. Vanhainkotihoitoon ja yli 65-vuotiaille kohdentuneen terveyskeskusten vuodeosastohoidon menoiksi voidaan arvioida siten noin 1,15 miljardia euroa vuonna 2001. Ikääntyneiden palvelujen menot ovat nousseet hieman viime vuosina. (SVT, 2003) Väestön ikääntyminen nostaa vanhusmenoista (palvelut + tulonsiirrot) erityisesti eläkemenoja. Parkkisen (Parkkinen, 2002) arvion mukaan vanhusmenot pystytään tulevaisuudessa kuitenkin rahoittamaan kohtuullisella talouskasvulla ja ikääntyneiden hoidon tarpeen siirtymisellä. Parkkinen on myös esittänyt, että lisääntyvien verotulojen takia vanhuusmenojen nettomenot eivät kasvaisikaan samassa määrin kuin bruttomenot (Parkkinen, 2004b). Menojen kasvua voidaan hillitä myös muuttamalla palvelurakennetta siten, että koti- ja asumispalvelujen osuutta lisätään. Lisäksi vanhusten itsenäistä selviytymistä tulee tukea (Vaarama & Voutilainen, 2002). Eläkepoliittisten ratkaisujen lisäksi vanhusmenoihin vaikuttaa myös

palvelujärjestelmän toimivuus eli esimerkiksi palvelujen organisointitapa, annettujen palvelujen laatu ja tuottavuus.

Henkilöstömäärän kasvu ja tuotettujen hoitopäivien määrän väheneminen vanhusten pitkäaikaishoidossa ovat osaltaan johtaneet siihen, että tuottavuuskehitys on viime vuosien aikana ollut negatiivinen (Lahtinen, 2003; Laine, 2003). Kuviossa 2 on esitetty vanhusten pitkäaikaisen laitoshoidon tuottavuuden muutos sekä paneeliaineistossa että vuosittaisissa poikkileikkausaineistoissa. Tuottavuus on laskenut noin seitsemän prosenttia vuodesta 2000 vuoteen 2003.² Viime vuosina tuottavuuden kehitys on ollut samansuuntainen koko terveydenhuollossa (ks. Hjerppe et al., 2003). Yleisesti terveydenhuoltoon on viime vuosina suunnattu aiempaa enemmän voimavaroja. Raha ei kuitenkaan muutu välttämättä suoraan tuotoksiksi, vaan osa lisärahoituksesta menee esimerkiksi hintoihin (palkkoihin) tai käytetään tehottomasti. Tuottavuus voi siten laskea, kun voimavaroja lisätään.

Edellä kuvatut vanhustenhuollon toimintaympäristön muutokset ovat muokanneet palvelurakennetta ja lisänneet vanhustenhuollon poliittista painoarvoa. Laitoshoidon järjestämisen ja vanhustenhuollon kansan- ja kuntataloudellisen merkityksen kannalta vanhustenhuoltoa on tulevaisuudessa arvioitava laajasti. Olennaisimpia kysymyksiä tulevaisuudessa ovatkin, mitä palveluihin käytetyillä voimavaroilla kyetään saamaan aikaan ja millaisia vaikutuksia toimintaympäristön muutoksella on.



KUVIO 2. Pitkäaikaisosastojen tuottavuuden kehitys vuosina 2000–2003

² Tuottavuus on laskettu jakamalla pitkäaikaishoito-osastojen kokonaiskustannukset asiakasrakenne- ja hoitopäivillä (RUG-III/22) hoitopäivillä. Kustannukset on deflatoitu julkisten menojen sosiaali- ja terveys-toimen hintaindeksillä ja vuosi 2000 on kiinnitetty arvoon 100. Aineisto muodostuu RAI-tietojärjestelmän käyttöönotto ja pitkäaikaishoidon benchmarking -hankkeessa mukana olevista osastoista (Noro, 2005; Noro et al., 2001).

3 Vanhusten laitoshoidon erityispiirteet taloustieteen näkökulmasta

3.1 Toimialan erityinen luonne

Sosiaali- ja terveydenhuollon toimintaympäristöön, erityisesti palvelujen tuotantoon ja markkinoihin liittyy erityispiirteitä, jotka vaikuttavat alan toimijoiden käyttäytymiseen ja markkinoiden toimintaan. Kenneth Arrow (Arrow, 1963) on esittänyt, miten terveydenhuollon markkinat eroavat ideaalimarkkinoista hyödykkeiden markkinakelpoisuuden, riskien ja epävarmuuden sekä kysynnän ja tarjonnan suhteen. Lisäksi terveystaloudessa palvelun käyttö ja palvelun tuotanto tapahtuvat samanaikaisesti, jolloin asiakas ei voi kokeilla tai arvioida hyödykettä ennen sen kuluttamista. Terveys- ja sosiaalipalvelujen tuotannon tutkiminen ja taloudellinen tarkastelu edellyttävät näiden erityispiirteiden huomioimista. Seuraavissa luvuissa käsitellään vanhusten laitoshoidon erityispiirteitä taloustieteen ja toimialan erityisyyden näkökulmasta. Luvuissa arvioidaan myös sitä, miten nämä erityispiirteet voivat vaikuttaa tämän tutkimuksen kohteina olevien palvelujen laatuun ja tuotannolliseen tehokkuuteen.

3.2 Tiedon epäsymmetria

Tiedon epäsymmetria tuottajan tai myyjän ja asiakkaan välillä koskee laajasti erilaisia tavaroita ja palveluja. Myyjällä onkin todennäköisesti lähes aina enemmän tietoa tuottamansa hyödykkeen laadusta kuin asiakkaalla. Erityisen ongelmallinen tiedon epäsymmetria on kuitenkin tilanteessa, jossa asiakkaan mahdollisuudet ja kyvyt arvioida laatua ovat heikentyneet. Suurimmalla osalla pitkäaikaisessa laitoshoidossa olevista asiakkaista on heikentynyt kognitiivinen toimintakyky. Noin 70 prosentilla vanhainkotien ja terveyskeskusten vuodeosastojen asiakkaista on vähintään keskivaikea kognition heikentyminen ja vain joka kymmenennellä asiakkaista ei ole minkäänasteista kognition heikentymää (SVT, 2003). Laitoshoidossa oleva asiakas ei siten välttämättä kykene ainakaan systemaattisesti ja yhtäjaksoisesti seuraamaan ja arvioimaan annetun hoidon laatua ja määrää. Asiakas ei myöskään välttämättä kykene vaatimaan laadukkaampaa hoitoa tai vaihtamaan palvelun tuottajaa. Tiedon epäsymmetrian vuoksi kysynnän (käytön) laatuujousto onkin laitoshoidossa todennäköisesti hyvin pieni. Ongelma on vielä merkittävämpi, jos asiakkaalla ei ole omaisia tai ystäviä, jotka säännöllisesti ottavat osaa hoitoon. Stakesissa vuonna

2002 tehdyn aikamittaustutkimuksen mukaan noin 50 prosentilla pitkäaikaishoidon asiakkaista ei ollut ketään ystävää tai sukulaista, joka olisi vieraillut asiakkaan luona viikon aikamittaaseurannan aikana (Laine, Noro, & Finne-Soveri, 2004).

Annetun hoidon laatua koskeva tiedon epäsymmetria asiakkaiden ja tuottajan välillä onkin ilmeinen ongelma vanhusten laitoshoidossa. Tilanne on todennäköisesti samankaltainen esimerkiksi vammaisten laitoshoidossa. Tiedon epäsymmetrian takia asiakkaan on tarkoituksenmukaista luoda palvelutuottajaan pitkäaikainen ja luottamuksellinen hoitosuhde, joka siten takaisi hänelle hyvän hoidon ja kohtelun kaikissa olosuhteissa (Le Grand et al., 1992). Niin kutsutun *contract failure* teorian mukaisesti kuluttajat tai asiakkaat ovat taipuvaisia valitsemaan voittoa tavoittelemattoman palvelutuottajan voittoa tavoittelevan sijaan, jos kuluttajan mielestä on etukäteen vaikea arvioida tarkasti annetun palvelun määrää tai laatua (ks. Hansmann, 1987). Epäsymmetrinen informaatio hoivapalveluissa merkitsee palvelun tuottajalle myös suurta moraalista vastuuta hyvän ja inhimillisen hoivan tarjoamisesta kaikille asiakkaille asiakkaan fyysisestä, psyykkisestä, kognitiivisesta tai sosiaalisesta toimintakyvystä riippumatta. Asiakkaan ja tuottajan välillä vallitseekin terveyspalveluissa erityinen luottamussuhde, jossa asiakas odottaa, että esimerkiksi lääkäri toimii hänen parhaakseen tai siten kuin asiakas toimisi itse, jos hänellä olisi käytettävänään kaikki päätöksenteossa tarvittava informaatio. Tätä suhdetta kuvataan päämies-agenttisuhteena.

3.3 Toiminnan laadullisten ulottuvuuksien arvioiminen

Palveluissa laatu saattaa vaikuttaa yksikön tuottavuuteen ja kustannuksiin, joten se tulisi ottaa huomioon myös taloudellisessa tarkastelussa. Terveystalouden tuottavuustutkimuksia on kritisoitu siitä, ettei niissä ole otettu laatua huomioon (Newhouse, 1994). Laatu on moniulotteinen käsite, jonka määrittelemisen ja mittaamisen erilaisissa hoivapalveluissa on vaikeaa (ks. Berg et al., 2002; Finne-Soveri, 1999). Laadukkaan palvelun kriteerit riippuvat siitä, kenen näkökulmasta laatua tarkastellaan eikä laadulle ei ole olemassa yhtä objektiivista mittaria (Ovretveit, 1998). Eri tahojen määritelmät ja kokemukset laadusta saattavat olla hyvinkin erilaisia. Vaikka laadusta käsitteellisellä tasolla päästäisiinkin jonkinlaiseen yhteisymmärrykseen, empiirisessä tutkimuksessa tuskin voidaan ottaa huomioon kaikkia laadun ulottuvuuksia. Laatua joudutaan tarkastelemaan vain jostakin näkökulmasta ja vain joillakin mittareilla.

Suomessa vanhusten hoidon valtakunnallisissa laatusuosituksissa (Vaarama et al., 2001) esitetään suuntaa-antavasti hyvän henkilöstömitoituksen määrää laitoshoidossa, joka on 0,8 hoitajaa yhtä asiakasta kohti. Lisäksi suositellaan, että kunnis-

sa olisi käytössä jokin laadunseurantajärjestelmä. Suositukset eivät kuitenkaan ole kuntia sitovia eivätkä juurikaan tätä yksityiskohtaisempia. Laaja-alainen laadunseuranta edellyttää seurantajärjestelmää ja systemaattisia arviointivälineitä, joita Suomessa ei kaikissa kunnissa ole. Suomessa laitos- ja sairaalahoidon tuottavien julkisten yksiköiden laadusta ei nykyisin ole saatavissa julkisesti vertailukelpoista tietoa kuten esimerkiksi Yhdysvalloissa tai Englannissa, jossa esimerkiksi julkiset sairaalat on luokiteltu internetissä tähtipisteytyksellä samaan tapaan kuten matkakesitteissä hotellit. Pisteytys perustuu useaan eri laatuindikaattoriin (ks. The Hospital Report..., 2001; Loeb, 2004). Laadun mittaamiseen liittyvien vaikeuksien ja puutteiden vuoksi hinta onkin todennäköisesti usein ainoa selvästi mitattava tekijä, josta Suomessa voidaan palvelujen järjestäjien ja tuottajien välillä sopia (Laine, 1999). Laadun mittaamisen ja arvioimisen vaikeus yhdessä epäsymmetrisen informaation kanssa muodostavat näin ollen monin tavoin ongelmallisen tilanteen vanhusten laitoshoidossa ja muissakin hoivapalveluissa.

3.4 Hintajärjestelmän poikkeavuus

Vanhusten pitkäaikainen laitoshoidon on Suomessa pääasiassa julkisen sektorin tuottamaa ja verovaroin rahoitettua, jolloin asiakkaat eivät suoranaisesti maksa oman hoitonsa kustannuksia. Tällöin palvelun hinta asiakkaalle ei kuvaa asiakkaan maksuhalukkuutta. Perinteisesti taloustieteessä asiakkaiden tai kuluttajien kysyntäkäyrän osoittama maksuhalukkuus ilmaisee rajahyödyn, mutta tämä ei päde julkisesti järjestetyssä laitoshoidossa. Laitoshoidon lopullinen kustannus asiakkaalle ei myöskään välttämättä määräydy ainoastaan annetun hoidon laadun tai määrän perusteella, vaan myös hallinnollisten kustannusten, tuottajalle annetun budjetin ja asiakkaan tuloihin perustuvan hoitopäivämaksun³ mukaisesti. Näin ollen julkisesta palvelusta asiakkaalle kohdistuva kustannus ei välttämättä ole mitenkään yhteydessä asiakkaan maksuhalukkuuteen, hoidosta saatavaan rajahyötyyn, annetun hoidon laatuun ja määrään. Edellä esitetyt erityispiirteet koskevat luonnollisesti monia muitakin julkisia palveluja.

³ Laki sosiaali- ja terveydenhuollon asiakasmaksuista 3.8.1992/734, § 7c: ”Pitkäaikaisessa laitoshoidossa olevalta peritään maksukyyn mukaan määräytyvä maksu. Maksu voi olla enintään 80 prosenttia hoidossa olevan tässä laissa tarkoitetuista kuukausituloista, jollei tässä laissa toisin säädetä. Maksu voidaan kuitenkin määrätä enintään sen suuruiseksi, että hoitoa saavan henkilön henkilökohtaiseen käyttöön jää kuukausittain vähintään 80 euroa. Jos pitkäaikaisessa laitoshoidossa oleva on välittömästi ennen laitoshoidon alkamista elänyt yhteistaloudessa avioliitossa tai avioliitonomaisissa olosuhteissa ja hänen kuukausitulonsa ovat suuremmat kuin puolison kuukausitulot, maksu määräytyy puolisojen yhteenlaskettujen kuukausitulojen perusteella. Pitkäaikaisessa laitoshoidossa olevalta perittävä maksu voi olla enintään 40 prosenttia edellä mainituin perustein yhteenlasketuista kuukausituloista. Laitoshoidossa olevan henkilökohtaiseen käyttöön tulee kuitenkin jäädä vähintään 80 euroa kuukaudessa. Jos molemmat tässä momentissa tarkoitettut puoliset ovat pitkäaikaisessa laitoshoidossa, maksu määräytyy kuitenkin siten kuin 1 momentissa säädetään.”

3.5 Asiakkaan rajoitettu valinnanmahdollisuus

Pääosa vanhusten pitkäaikaisesta laitoshoidosta tuotetaan julkisissa laitoksissa. Vuonna 2002 vanhainkotihoidosta 89 prosenttia oli kuntien, 11 prosenttia järjestöjen ja alle 1 prosentti yritysten tuottamaa (SVT, 2003). Terveyskeskusten vuodeosastohoitoa tuotettiin käytännössä ainoastaan kuntien ja kuntayhtymien terveyskeskussairaaloissa. Pitkäaikaishoidon palvelurakenne vaihtelee jonkin verran kunnittain. Suurissa kunnissa vanhusten laitoshoidoa tuotetaan tavallisesti palvelutaloissa, vanhainkodeissa ja terveyskeskussairaaloissa, mutta pienissä kunnissa pitkäaikaishoitoa voidaan tarjota joko pelkästään kunnan vanhainkodissa tai terveyskeskuksessa. Käytännössä tämä merkitsee, että asiakkaan mahdollisuus valita hoitopaikkansa voi olla rajattua, erityisesti pienissä kunnissa. Toisaalta myös suurissa kaupungeissa asiakkaan valinnanmahdollisuudet voivat olla rajattuja, jos asiakkaan hoitopaikka määräytyy ensisijaisesti asuinalueen, esimerkiksi kaupungin suurpiirin mukaan tai niin sanotun SAS-toiminnan kautta (Selvitä, Arvioi, Sijoi-ta). Kunnan ja laitoksen sisällä asiakas voi periaatteessa vaihtaa hoitopaikkaa, jos vapaita paikkoja on tarjolla, mutta pitkäaikaishoitopäätöksen saanut asiakas ei voi toistaiseksi (vuonna 2004) lain mukaan muuttaa toisen kunnan ylläpitämään vanhainkotiin tai terveyskeskukseen ilman vastaanottavan kunnan suostumusta. Lisäksi asiakkaan kognition heikentymä voi osaltaan rajoittaa asiakkaan päätöksen-tekokykyä ja kykyä vaikuttaa hoitopaikkaansa.

3.6 Budjettirajoitteet ja tulonmuodostus

Julkisen pitkäaikaishoitoyksikön budjetti määritellään etukäteen kunnan ja toimialasektorin taloudellisen tilanteen ja arvioitujen asiakasmaksutuottojen perusteella. Määritelty budjetti toimii siten budjettirajoitteena palvelutuotannolle. Julkisel-la organisaatiolla ei ole usein myytäviä hyödykkeitä, joilla tulonmuodostusta voi lyhyellä aikavälillä lisätä. Näin ollen laitoksen tai yksittäisen osaston mahdollisuus vaikuttaa tulonmuodostukseen lyhyellä aikavälillä on rajattu. Tällaisessa toimintaympäristössä on mahdollista, että palvelun tuottaja pyrkii pysymään annetussa budjetissa toiminnan laadullisella muutoksella, esimerkiksi vaikuttamalla työvoiman käyttöön, määrään ja laatuun.

3.7 Ihmistyön korvaamattomuus

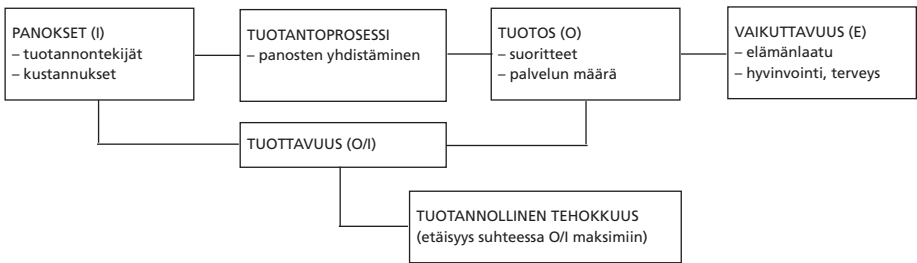
Henkilöstö on vanhustenhuollon tärkein ja kallein voimavara. Pitkäaikaishoitoa tuottavissa laitoksissa noin 70–80 prosenttia kustannuksista on henkilöstökustannuksia. Yksikkötasolla hyvin merkittäviä kustannussäästöjä voidaan näin ollen saavuttaa vain työvoiman käyttöön liittyvillä muutoksilla, joita ovat esimerkiksi henkilöstön määrään, sijaisten käyttöön ja työvuoroihin liittyvät järjestelyt. Työvoimaa ei voida juurikaan korvata muilla tuotannontekijöillä, esimerkiksi pääomalla. Hoi-va-alan työvoimavaltaisuuden vuoksi uudella teknologialla tai uusilla hoitomenetelmillä ei voida saavuttaa merkittäviä kustannussäästöjä ja tuottavuuden parantumista. Toisaalta ei todennäköisesti ole myöskään odotettavissa kustannusten merkittävää kohoamista teknologisen kehityksen johdosta, kuten muualla terveydenhuollossa. Useissa tutkimuksissa on osoitettu, että vanhustenhoidossa henkilöstötekijät ovat yhteydessä toiminnan laatuun (Dellefield, 2000; Harrington, Zimmermann, Karon, Robinson, & Beutel, 2000; Syvänen, 2003). Laitoksen tai osaston pyrkimys parantaa toiminnan laatua saattaa edellyttää lisäpanostusta esimerkiksi henkilöstön koulutustasoon ja lukumäärään. Tämä taas nostaa todennäköisesti palkkakustannuksia, mikä puolestaan voi johtaa tuottavuuden heikentymiseen.

4 Palvelutuotantoprosessi, tuotannollinen tehokkuus ja laatu

4.1 Terveyspalvelujen tuotantoprosessi ja taloustieteellinen tuottavuustutkimus

Tämän tutkimuksen empiirisenä tutkimuskohteena on vanhusten pitkäaikainen laitoshoito, jonka tuotantoa ja tutkimusta voidaan hahmottaa prosessinäkökulmasta kuvion 3 mukaisesti. Tuotantoyksiköiden käytettävissä on rajoitettu määrä voimavaroja (panoksia), joita yhdistelemällä (tuotantoprosessi) tuotetaan tuotoksia, joilla vuorostaan on tiettyjä vaikutuksia (vaikuttavuus).

Tässä tutkimuksessa palvelutuotantoprosessia tarkastellaan taloustieteellisen tuottavuustutkimuksen näkökulmasta. Tutkimuksessa arvioidaan erityisesti tuotosten ja panosten suhdetta (O/I). Kuviossa 3 on huomattava, että tuotannollinen tehokkuus on tuottavuuteen eikä vaikuttavuuteen (E) tai kustannus-vaikuttavuuteen (E/I) liittyvä käsite, kuten tehokkuudella terveystaloustieteessä usein tarkoitetaan. Väärinymmärryksen välttämiseksi tässä tutkimuksessa käytetään tuottavuuteen liittyvänä tehokkuuskäsitteenä tuotannollista tehokkuutta.⁴ Siitä huolimatta,



KUVIO 3. Terveystuotantotoiminnan elementit ja tehokkuustutkimuksen osatekijät

⁴ Barr (Barr, 1998) käyttää tuotannollisesta tehokkuudesta termiä *productive efficiency*. Ylipäättään tehokkuuden (efficiency) käsite on ongelmallinen siksi, että sillä on erilainen merkitys esimerkiksi tuotantoteoriassa (Shepard, 1970) ja terveystaloustieteessä (Drummond, O'Brien, Stoddart, & Tarrance, 1997; Sintonen, Pekurinen, Jämsen, Linnakko, & Vinni, 1987). Tuotantoteoriassa tehokkuus määritellään eksaktisti. Tehokkuudella tarkoitetaan yksikön etiäisyyttä ei-dominoitujen yksikköjen muodostamasta tehokkuusrintamasta (eli tehokkaan tuotannon rintamasta). Knapp (Knapp, 1984) on esitellyt osuvasti tehokkuuskäsitteen eri merkityksiä ja sovellustapoja terveystaloustieteessä. Yksi merkittävä erottelu liittyy siihen, kuvataanko tehokkuudella suoritteiden aikaansaamista (outputs) eli itse asiassa tuottavuutta vai lopputuloksia (effectiveness, outcomes) eli vaikuttavuutta. Molemmissa tapauksissa olisi Knappin mukaan kuitenkin hyväksyttävää käyttää tehokkuustermiä. On myös huomattava, että terveystaloustieteen sisällä eri maissa tehokkuus (efficiency) ja tuottavuus (productivity) voivat olla lähes synonyymeja tai kuvion 3 mukaisesti täysin eri asioita (ks. Eggleston, Grossman, & Cutler, 2004; Linna, 1999; Sintonen et al., 1987). Terveystaloustieteessä tehokkuuden arvioinnillekin (efficiency evaluation) on muodostunut erityinen merkityksensä. Sillä tarkoitetaan yleensä taloudellista arviointia (economic evaluation) eli kustannusvaikuttavuus- ja kustannushyöty- ja kustannusutiliteettianalyysia (ks. Drummond et al., 1997; Sefton, Byford, McDaid, Hills, & Knapp, 2002; Sintonen et al., 1987). Näissä hyötyä mitataan usein rahamäärässä ja vaikuttavuutta esimerkiksi laatu painotteisina elinvuosina (QALYs). Näillä taloudellisen arvioinnin menetelmillä voidaan vastata kaikkiin palvelutuotannon peruskysymyksiin, koska on mitattu sekä käytetyt voimavarat että vaikutukset.

että tuottavuus ei sinänsä kerro mitään palvelujen vaikuttavuudesta, oletetaan tämän tutkimuksen tuottavuustarkastelussa, että laitoshoito on sinänsä asiakkaille tarpeellisia ja hyödyllistä. Mikäli käytössä olisi sekä käyttökelpoisia vaikuttavuusmittareita että vaikuttavuustutkimuksen mahdollistava aikasarja-aineisto, mikään ei estäisi laajentamasta tarkastelua tuotannollisesta tehokkuudesta myös siihen, millä kustannuksilla eri yksiköt tuottavat laatupainotteisia elinvuosia, kun yksiköiden väliset erot asiakasrakenteessa on otettu huomioon.

Tuottavuus voi parantua teknologisen kehityksen tai toimintaympäristössä tapahtuvien muutosten myötä. Teknologisella kehityksellä tarkoitetaan esimerkiksi aiempaa parempia hoitomenetelmiä, joiden avulla tehokkaan tuotannon rintamaa voidaan ikää kuin työntää eteenpäin. Toimintaympäristömuutoksilla tarkoitetaan puolestaan esimerkiksi kilpailun lisääntymistä ja palvelurakenteen muuttumista. Tuotannollisella tehokkuudella tarkoitetaan siis jonkin yksikön tuottavuutta suhteessa määriteltyn optimitasoon eli parhaimpaan mahdolliseen tuottavuuteen. Tuotannollinen tehottomuus kuvaa sitä, missä määrin tehottomat yksiköt tuullaavat voimavaroja suhteessa tuotannollisesti tehokkaisiin yksiköihin. Tuottavuus voikin parantua myös siten, että tuotannollisesti tehottomien yksiköiden etäisyys tehokkuusrintamaan pienenee.

Tässä tutkimuksessa tuotannollisen tehokkuuden alakäsitteitä ovat esimerkiksi tekninen tehokkuus ja kustannustehokkuus. Yksikkö on teknisesti tehokas suhteessa toiseen yksikköön, jos se pystyy tuottamaan saman tuotosmäärän pienemmällä panosmäärällä tai jos se samalla panosmäärällä saa aikaan enemmän tuotosta (Farrell, 1957; Koopmans, 1951). Teknisesti tehokas yksikkö ei tällöin voi tuottaa enempää tuotosta annetuilla voimavaroillaan ja sen hetkiselä tuotantoteknologian tasolla. Kustannustehokkuudella tarkoitetaan tuotoksen aikaansaamisesta mahdollisimman alhaisin kustannuksin. Sen laskeminen edellyttää tietoa muun muassa panoshinnoista. Kustannustehokkuuden saavuttamiseksi yksikön on oltava sekä teknisesti tehokas että käytettävä voimavarojaan eli eri panoksia optimaalisesti (ks. Coelli, Rao, & Battese, 1998).

Tuotantoteoriassa panosten ja tuotosten suhdetta kuvataan tavallisesti tuotantofunktiolla, joka kuvaa eri panosmäärillä ja tietyllä teknologian tasolla saatutettavan maksimaalisen tuotannon määrä. Perinteisesti tuotantofunktion tuotantontekijöihin sisällytetään työ, pääoma, koneet ja laitteet. Tuotantoteoriassa tuotantofunktion estimoinnille läheinen lähestymistapa on kustannusfunktion estimointi, jolla voidaan esittää panoshintojen ja tuotosten yhteys kokonaiskustannuksiin (ks. Shepard, 1970). Tuotantoteoriassa eri hyödykkeiden laadulliset ominaisuudet huomioidaan tavallisesti siten, että eri laatuiset hyödykkeet ovat eri tuotoksia, joilla on eri markkinahinta. Perinteistä tuotantoteoreettista lähestymistapaa on terveystaloustieteessä pyritty laajentamaan sisällyttämällä tuotantofunktion myös tuotantoprosessin aineettomia voimavaroja, joita ovat esimerkiksi henkilökunnan työilmapiiri ja työympäristö (Knapp, 1984; Knapp & Davies,

1981). Tuotantoteoriaan liittyvä kritiikki kertoo siitä, että sosiaali- ja terveyspalvelujen tuotantoprosessin käsitteellistäminen perinteisten tuotantoteoreettisten mallien avulla ei välttämättä kuvaa terveydenhuoltotoimintaa riittävän kattavasti ja että kulloinkin tarkasteltavaa tuotosta ei ole välttämättä edes mahdollista jakaa sen eri laadullisten ominaisuuksien (hyvä tai huono) mukaan. Ilmiön laadulliset ulottuvuudet tulisi ottaa tarkasteluun mukaan muulla tavoin. Kritiikissä on kyse myös siitä, että mittarit eivät ole erityyppisissä hoivapalveluissa ole vielä kehittyneet siten, että laaja-alainen toiminnan kustannukset ja vaikuttavuuden sisältävä arviointi olisi ollut laajamittaisesti mahdollista.

Terveyshalloustieteessä palvelutuotannon lopullisena tavoitteena pidetään siis hoidon tai palvelun aikaansaamaa muutosta ihmisen terveydentilassa, hyvinvoinnissa ja elämänlaadussa (Drummond et al., 1997). Muutoksen ei kuitenkaan tarvitse välttämättä merkitä terveydentilan parantumista, vaan hoidon vaikuttavuus voi tarkoittaa myös terveydentilan ylläpitämistä tai huononemisen hidastumista verrattuna tilanteeseen ilman hoitoa tai verrattuna johonkin toiseen hoitoon. Vanhusten pitkäaikaisen laitoshoidon tavoitteena ei useinkaan voi olla terveydentilan parantaminen, vaan pikemminkin toimintakyvyn ja terveydentilan heikkenemisen hidastaminen ja ylipäätään inhimillinen hoiva. Tämän tutkimuksen tutkimusasetelma ei suoranaisesti mahdollista terveyshalloustieteellistä kustannus-vaikuttavuuden arviointia. Eräät tutkimuksessa käytetyt kliiniset laatuindikaattorit kuvaavat kuitenkin hoidon lopputulosta eli vaikuttavuutta. Tästä syystä välisuoritteita ja laatuindikaattoreita voidaan siten käyttää epäsuorasti lopputuotosten eli vaikuttavuuden indikaattoreina (Knapp, 1984; Sintonen et al., 1987).

Tässä tutkimuksessa tarkastellaan tuotannollista tehokkuutta ja hoidon laatua yksinomaan palvelutuottajan ja palvelutuotannon näkökulmasta. Asiakkaan terveyden kannalta tarjottu palvelu on vain yksi osa laajempaa kokonaisuutta. Itse palvelun lisäksi asiakkaan terveyteen ja hyvinvointiin vaikuttavat myös muun muassa yksilön ominaisuudet ja ympäristötekijät (Evans & Stoddart, 1990; Grossman, 1972), joita kaikkia ei tässä tutkimuksessa voida mitata tai ole tarkoituksenmukaista ottaa huomioon.

4.2 Laadun käsite

Hoidon laadun arviointi ja kiinnostus laatua kohtaan eivät ole sosiaali- ja terveydenhuollossa mitenkään uusia ilmiöitä. Jo 1800-luvulla ja 1900-luvun alussa pyrittiin sekä Englannissa että Yhdysvalloissa arvioimaan sairaaloiden hoidon laatua (Mainz, 2004). Suomessa sosiaali- ja terveydenhuollon laatuun liittyvien tutkimusten ja selvitysten määrä on merkittävä (Makkonen & Inkinen, 2001). Myös aivan viime vuosina palvelujen laatu on ollut esillä julkisessa keskustelussa. Vuonna 2000 käytiin tiedotusvälineissä vilkas keskustelu vanhusten laitoshoidon laadusta, min-

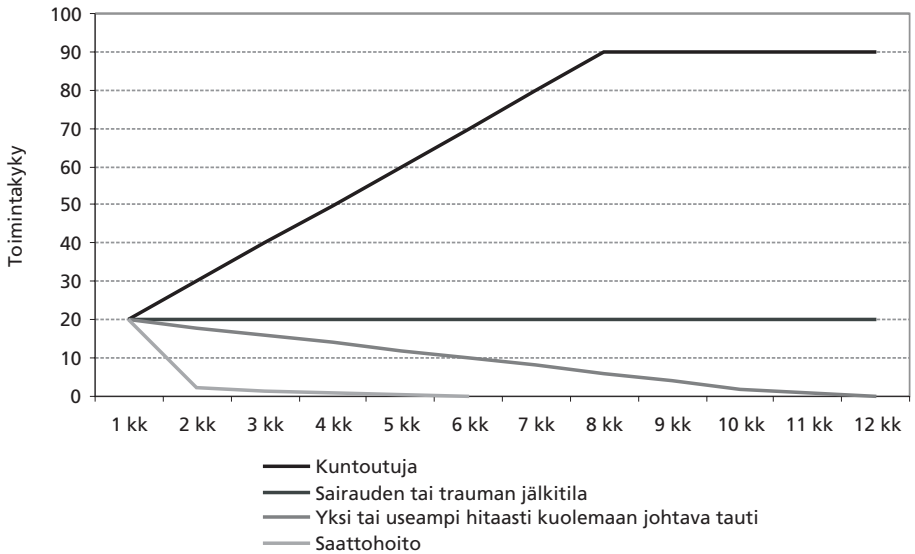
kä seurauksena Stakes, Suomen Kuntaliitto ja Sosiaali- ja terveysministeriö laativat valtakunnalliset suositukset vanhusten hoidon ja palvelujen laadun kehittämiseksi (Vaarama et al., 2001).

Laatu-käsitteen sisällöstä ja sen mittaamisesta ei kuitenkaan ole saavutettu yksimielisyyttä. Laadun määritelmiä on lukuisia ja eri tieteenaloilla puhutaan vaihtelevasti laadusta tai vaikuttavuudesta, vaikka niillä käsitteellisesti tarkoitettaisiinkin samaa asiaa. Ovretveit (Ovretveit, 1998) on määritellyt laadun käsitteen hyvin yleisesti ja kattavasti. Hänen määritelmässään yhdistyvät hoidon laatu, oikeudenmukaisuus (jakauma) ja tehokkuus kustannus-vaikuttavuus mielessä.

Health service quality is meeting the health needs of those most in need at the lowest cost, and within regulations.

Vanhusten laitoshoidon eräät erityispiirteet vaikuttavat siihen, miten palvelun laatu tulisi tällaisissa palveluissa ymmärtää. Kanen (Kane, 1998) mukaan pitkäaikaishoidossa olevien asiakkaiden hoidon tavoitteet eivät ensinnäkään ole yhtäläisiä. Asiakkaan voinnin merkittävä paraneminen, kotiutuminen tai siirtyminen laitoksesta kotiin ei ole mahdollista kuin harvoille pitkäaikaishoitoasiakkaille. Suurin osa asiakkaita kärsii joko kroonisista fysiologisista sairauksista tai dementiasta (SVT, 2003), jotka pitkällä aikavälillä heikentävät terveydentilaa. Lisäksi osa pitkäaikaishoidossa olevista asiakkaita on terminaalivaiheessa tai lähes tiedottomassa tilassa (SVT, 2003). Kuviossa 4 on esitetty erityyppisten asiakkaiden toimintakyvyn kehitys laitoshoidossa. Pystyakselilla on kuvattu fyysisen, psyykkisen, kognitiivisen tai sosiaalisen toimintakyvyn taso (100 = täysin itsenäinen, 0 = täysin toisen ihmisen varassa oleva) ja vaaka-akselilla aika.

Kanen (Kane, 1998) mukaan hyvä laatu tällaisessa toimintaympäristössä voikin tarkoittaa sitä, että asiakas voi paremmin kuin olisi odotettavaa muualla tai tilanteessa ilman hoitoa. Pitkäaikaishoidon yleisenä tavoitteena on pyrkimys sekä asiakkaiden maksimaaliseen toimintakykyyn että kärsimysten ja kipujen lievittämiseen (Finne-Soveri, 1999, 2005). Näin ollen laitoshoidon laatua voitaisiin esimerkiksi arvioida sillä, miten hyvin asiakkaan fyysisen toimintakyvyn laskua onnistutaan hidastamaan (Finne-Soveri, 1999; Noro et al., 2001). Mikäli hoidon laatu ymmärretään tällä tavoin, tarkoittaisi se käsitteellisesti samaa kuin hoidon vaikuttavuus (vrt. kuvio 3, s. 29).



Lähde: Finne-Soveri 2001 (teoksessa Noro et al. 2001)

KUVIO 4. Asiakkaan fyysisen, psyykkisen, kognitiivisen tai sosiaalisen toimintakyvyn kehitys laitoshoidossa

4.3 Laadun ulottuvuudet

Laadun määrittely riippuu merkittävästi siitä, kenen näkökulmasta määritelmä tehdään. Hoitotyöntekijöiden ja muiden ammattilaisten näkökulmasta laatua voidaan arvioida sen perusteella, että saavatko asiakkaat tarvitsemaansa hoitoa tai onko käytettävissä oleva henkilöstö oikeanlaista ja tekeekö se oikeita asioita asiakkaiden tarpeiden tyydyttämiseksi. Palvelujen järjestäjän näkökulmasta toiminnan laatua voidaan arvioida siten, että pystytäänkö asiakkaiden tarpeisiin vastaamaan käyttäen voimavaroja mahdollisimman optimaalisesti ja ottaen huomioon palvelutuotantoa rajoittavat ja ohjaavat säädökset. Asiakkaan näkökulmasta laatua arvioidaan henkilökohtaisen kokemuksen kautta ja sen mukaan, miten hyvin palvelu vastaan asiakkaan omia haluja ja tarpeita (Øvretveit, 1998). Yllä olevista määritelmistä voidaan havaita, että laadulla voidaan tarkoittaa myös jotakin muuta kuin vaikuttavuutta ja hoidon lopputuotoksia.

Laadun ulottuvuudet määritellään usein Donabedianin (Donabedian, 1988) klassisen määritelmän mukaisesti. Sen mukaan laadun kolme ulottuvuutta ovat rakenne- tai voimavaralaatu, prosessilaatu ja tulostaatu. Rakennelaatu kuvaa tuotantoprosessissa käytettävien voimavarojen ominaisuuksia, joita ovat esimerkiksi

henkilöstön ammattitaito, henkilöstön koulutusmahdollisuudet ja toimintaympäristön fyysiset puitteet. Prosessilaatu liittyy puolestaan toimintaprosesseihin, jotka johtavat toivottuihin tai ei-toivottuihin tuloksiin. Prosessilaatua ovat muun muassa hoitokäytännöt (esimerkiksi kuntotuttava työote), eri ammattiryhmien yhteistyön toimivuus ja asiakkaiden oikea sijoittelu laitoksessa. Tuloslaatu kuvaa tuotantoprosessin lopputuotosta, joita ovat esimerkiksi asiakastyytyväisyys, asiakkaiden terveydentila, ei haitalliset hoitotulokset ja voimavarojen optimaalinen käyttö. Taulukossa 1 on esitetty Øvretveitin (Øvretveit, 1998) mukaan laadun arvioinnin malli laadun olottuvuuksien ja arvioijan näkökulman mukaan.

Øvretveitin malli on analoginen terveydenhuoltotoiminnan mallin kanssa (ks. kuvio 3, s. 29). Molemmissa eri osatekijöitä voidaan tarkastella erikseen tai prosessina voimavaroista tuloksiin. Øvretveitin mallissa eri olottuvuudet voidaankin nähdä itseisarvona sinänsä. Esimerkiksi vuodeosaston miellyttävät ja korkealaatuiset tilat ja välineet voivat olla tärkeitä asiakkaille, vaikka niillä ei olisikaan yhteyttä annetun hoidon tuloksiin. Asiakkaat voivat myös olla valmiita maksamaan korkealaatuisesta ja viihtyisästä miljööstä. Mitä enemmän asiakkaat kokevat saavansa hyötyä ympäristöstä, sitä korkeampi on todennäköisesti heidän maksuhalukkuutensa. Maksuhalukkuutta voidaan arvioida myös palvelujärjestelmän näkökulmasta. Kuinka paljon yhteiskunta on valmis panostamaan esimerkiksi vanhainkotien toimitiloihin ilman, että on tietoa sen yhteydestä hoidon tuloksiin? Øvretveitin

TAULUKKO 1. Laadun arvioinnin malli laadun olottuvuuksien ja arvioijan mukaan

	Voimavaralaatu	Prosessilaatu	Tuloslaatu
Asiakas	Osaava ja kokenut henkilöstö, puhtaat ja miellyttävät tilat ja välineet	Kohtelias ja ystävällinen kohtelu, riittävä tiedon saanti, ei tarpeetonta kipua, hoidon nopea saatavuus tarvittaessa	Asiakastyytyväisyys, kivun lievitys, toimintakyvyn parantuminen, säädyllyinen kuolema
Työntekijät ja ammattilaiset	Hyvin koulutetut ammatinharjoittajat, oikeanlaiset asiakkaat ohjataan hoitoon, asiakkaista riittävä informaatio saatavilla, oikeanlaiset välineet, tukipalvelujen saatavuus	Oikeat diagnoosit, oikeat interventiot, valitusmenettely, työjärjestyksen noudattaminen, nopeat tukipalvelut, eri ammattiryhmien välinen kommunikatio	Hyvä terveys, ei haitallisia hoitotuloksia
Palvelujen järjestäjä	Riittävät voimavarat, hyvät ulkopuoliset palvelut ja tiedot	Ei tuhlausta, viiveitä eikä viivytyksiä, säädösten ja ohjeiden noudattaminen	Voimavarojen tehokas käyttö

Lähde: Øvretveit 1998 (suom., muokattu)

mallin mukaan laatua voidaan arvioida prosessina, jossa voimavara- ja prosessiladulla saadaan aikaan tuloslautua. Tällöin laadun arvioinnin painopiste on selvästi tuloslautun mittaamisessa.

4.4 Laadun ulottuvuuksien mittaaminen

Laadun ulottuvuuksia voidaan tutkia ja mitata eri näkökulmista joko subjektiivisten arvioiden tai objektiivisten mittausten avulla. Laadua voidaan arvioida kysymällä asiakkaiden, henkilökunnan ja palvelun järjestäjän kokemuksia ja mielipiteitä. Tähän voidaan käyttää esimerkiksi asiakaskyselyä. Objektiivinen mittaus edellyttää puolestaan laadun käsitteen ja sen empiiristen vastineiden määrittelmistä ennalta. Tällöin on oltava käsitys siitä, millaista on laadukas hoito ja miten sitä voidaan mitata. Siten laadun mittaaminen ei itse asiassa voi koskaan olla täysin objektiivista, koska objektiivisenkin mittaus perustuu jonkun tahon tekemisiin määritelmiin ja ennako-oletuksiin. Objektiivisuutta tavoittelevan laadun mittamisen keskeisinä instrumentteina ovat systemaattiset indikaattorit, joilla todellista laatua pyritään mittaamaan. Yhtenä esimerkkinä systemaattisesta laadun mittamisesta ovat Zimmermannin johdolla Wisconsinin yliopistossa kehitetyt vanhusten pitkäaikaishoidon kliiniset laatuindikaattorit (Karon & Zimmermann, 1996; Zimmermann, Karon, & Arling, 1995), jotka perustuvat MDS 2.0-lomakkeen tietopohjaan. MDS 2.0 on pitkäaikaishoitoasiakkaan arviointiin käytettävä tietopohja, joka sisältää yli 300 muuttujaa ja 18 asiakkaan ominaisuuksia kuvaavaa osaluetta. Lomakkeen tietojen avulla voidaan laskea muun muassa Zimmermannin kehittämät kliiniset laatuindikaattorit, joita on kaiken kaikkiaan 41. Niiden avulla voidaan arvioida ei-toivottujen hoitoprosessien ja -tulosten yleisyyttä tiettyinä hetkenä ja tietyllä aikavälillä ilmenneitä uusia tapauksia. Tällaisia laatuindikaattoreita ovat esimerkiksi painehaavat, kaatuilu, liikkumista estävien välineiden käyttö ja toimintakyvyn aleneminen. Nämä laatuindikaattorit ovat laajasti käytössä Yhdysvalloissa, jossa laitosten laatuindikaattorit ovat myös julkisesti nähtävissä internetissä (The Hospital Report..., 2001; Loeb, 2004). Kliinisiä laatuindikaattoreita käytetään myös Suomessa (ks. Finne-Soveri, 2005).

Yleisin ja helpoin tapa toteuttaa laadun mittaus on kuitenkin asiakaskysely (Øvretveit, 1998). Pitkäaikaishoidossa asiakailta ja omaisilta voidaan kysyä kokemuksia annetun palvelun laadusta, henkilöstöstä ja toimitiloista. Asiakaskyselyihin liittyy yleisesti ottaen kuitenkin eräitä vakavia ongelmia. Ensinnäkin mittarit eivät ole välttämättä standardoituja eivätkä suoraan vertailukelpoisia eri yksiköiden välillä. Toiseksi asiakaskyselyt voivat olla herkkiä otoskoolle ja kadolle. Kolmanneksi asiakaskysely pitää suunnitella huolella, jotta kysymykset todella mittaavat hoidon laatua (Loeb, 2004). Erityyppisissä hoivapalveluissa asiakaskyselyihin liittyy vielä kaksi erityistä ongelmaa. Vanhusten laitoshoidossa asiakaskyselyt eivät välttämättä

tuota luotettavaa tietoa, jos asiakkaat tai omaiset uskovat, että vastauksilla voi olla haitallista vaikutusta annettuun hoitoon. Lisäksi asiakkaat saattavat olla kognitioltaan niin huonokuntoisia, että laadun systemaattinen ja ajallinen arvioiminen on mahdotonta. On kuitenkin syytä korostaa, että luonnollisesti myös objektiivisiin laadun mittareihin sisältyy riskejä ja sudenkuoppia. Ensinnäkin erityyppisille osastoille ja laitoksille voi systemaattisesti valikoitua tietyntyyppisiä asiakkaita, mikä saattaa aiheuttaa harhaa laatuvertailuissa. Jonkin haitallista hoidon tulosta kuvaavan ilmiön yleisyys ei välttämättä kerro laadusta ilman, että huomioidaan myös kyseisen ilmiön uudet tapaukset jollakin tietyllä ajanjaksolla. Toiseksi henkilöstön mittaama tai arvioima hoidon laatu voi heikentyä vain siitä syystä, että henkilöstö oppii koulutuksen myötä aiempaa luotettavammin arvioimaan hoidon laatua ja tunnistamaan haitalliset hoitoprosessit ja -tulokset (Karon & Zimmermann, 1996; Noro et al., 2001; Zimmermann et al., 1995). Käytännössä on mahdollista, että subjektiivisten ja objektiivisten laatumittarien antamat tulokset eivät ole yhdenmukaisia. Omaiset saattavat arvioida laadun heikommaksi kuin hoidossa olevat asiakkaat, mutta molemmat ryhmät voivat olla suhteellisen tyytyväisiä, vaikka kliiniset laatuindikaattorit osoittaisivat laatuongelmia. Koska laatu on moniulotteinen käsite, on myös mahdollista, että yksikkö voi toimia joillakin laadun osa-alueilla hyvin ja jollakin toisella taas heikommin.

Laadun käsitteellistämiseen ja mittaamiseen on siis olemassa useita mahdollisuuksia, joissa periaatteessa samaa ilmiötä (laatua) tarkastellaan toisiaan täydentävistä näkökulmista. Eräissä tuotannollista tehokkuutta käsittelevissä tutkimuksissa (ks. luku 6.2) toiminnan laatua on mitattu käyttämällä voimavaralaatua tai kliinisiä prosessi- ja tulostaatua kuvaavia indikaattoreita kuten esimerkiksi yksikön henkilöstön ammattirakennetta tai painehaavojen esiintyvyyttä. Laatumuuttujien valintakriteerien kuvaus ja ylipäättään niiden sisällöllinen arviointi on kuitenkin ollut aiemmissa tuotannollista tehokkuutta käsittelevissä tutkimuksissa suhteellisen niukkaa. Lisäksi tavallisimmin havaintoyksikkönä on ollut laitos, mikä luonnollisesti peittää alleen mahdollisesti merkittävänkin laitoksen osastojen välisen vaihtelun hoidon laadussa. On kuitenkin selvää, että empiirisessä tutkimuksessa joudutaan aina rajoittumaan tiettyyn laatuindikaattoriin ja tiettyihin indikaattoreihin. Laatu on moniulotteinen ilmiö eikä sitä voida operationalisoida yhdeksi mittariksi. Tässä tutkimuksessa laatua tarkastellaan Ovretveihin mallin mukaisesti (ks. taulukko 1, s. 34) ensisijaisesti ammattilaisten näkökulmasta määrittävänä prosessi- ja tulostaatuna. Tällöin laatu operationalisoidaan osastotasolla esiintyviksi haitalliseksi kliiniseksi hoitoprosesseiksi ja -tuloksiksi (Zimmermann et al., 1995). Lisäksi käytetään eräitä osaston voimavaralaatua kuvaavia muuttujia (ks. luku 8.9).

4.5 Palvelutuottajien tavoitteet ja kannustimet

Tämän tutkimuksen näkökulma on tuottajanäkökulma. Tässä luvussa tarkastellaan yksityisten ja julkisten (yleisemmin voittoa tavoittelemattomien) palvelutuottajien tavoitteita ja tuotannollisesti tehokkaaseen ja laadukkaaseen palvelutuotantoon kannustavia tekijöitä. Yksittäisten palvelutuottajien näkökulmasta tuotannon tavoitteet kiteytyvät usein tuotannolliseen tehokkuuteen eli siihen, kuinka paljon käytettävissä olevilla voimavaroilla onnistutaan tuottamaan tarjottavaa palvelua tai erilaisia aineellisia hyödykkeitä. Olennainen kysymys julkisten laitoshoidon tuottavien yksiköiden kohdalla onkin se, mikä kannustaa yksiköitä toimimaan tuotannollisesti tehokkaasti ja laadukkaasti ja miten toimintaympäristön erityisyys (ks. luku 3) vaikuttaa tavoitteisiin?

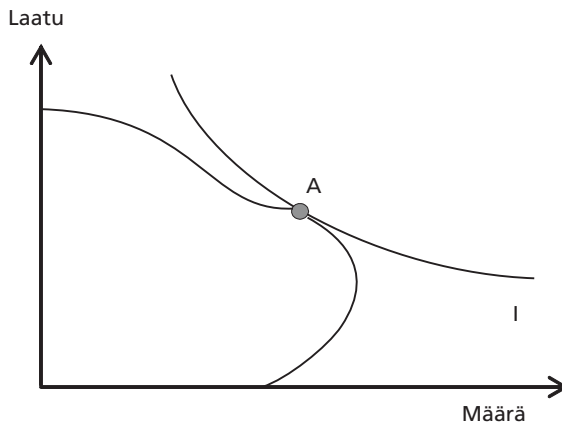
Täydellisen kilpailun markkinoilla yrityksen toiminta perustuu voiton maksimointiin tai vastaavasti kustannusten minimointiin. Mahdollisuus pitää liikevoittoa itsellään ja ylipäättään liiketoiminnallisten edellytysten ylläpitäminen toimivat kannusteina tuotannollisesti tehokkaalle palvelutuotannolle (Fizel & Nunnikho-ven, 1992). Koska julkiset yksiköt eivät toimi liiketaloudellisilla perusteilla, ovat niiden tavoitteetkin todennäköisesti erilaiset, jolloin myös kannustejärjestelmä toimii eri tavoin kuin yrityksissä. Terveystaloustieteessä on jo pitkään arvioitu esimerkiksi sairaaloiden ja lääkäreiden käyttäytymistavoitteita ja palvelujen tuottajan ja asiakkaan välistä erityistä suhdetta (esim. Arrow, 1963). Sairaaloita on käsitelty organisaationa, jolla on erilaisia tavoitteita kuten esimerkiksi tuotetun määrän ja hoidon laadun maksimointi (Newhouse, 1970) tai rahavirran maksimointi (Davis, 1972). Niskasen (Niskanen, 1968) mukaan julkiset organisaatiot pyrkivät puolestaan ennen kaikkea budjettinsa maksimointiin, mikä heijastaa julkisen sektorin toimimattomuutta (public sector failure). Sairaalaorganisaatiota on tarkasteltu myös toimijana, joka reagoi toimintaympäristön ja kannustinjärjestelmän muutoksiin (Ching-To, 1994; Propper, 1996). McGuiren (A. McGuire, 1985) mukaan sairaalaorganisaatioiden käyttäytymismotiiveihin ei kuitenkaan vaikuta mikään yksittäinen tekijä, vaan tekijöitä on tosiasiaa useita. Hänen mukaan sairaalaorganisaatioiden toimintaa voidaan selittää sekä sairaalan sisäisillä eli päätöksentekijöiden ominaisuuksiin liittyvillä tekijöillä että markkinaolosuhteisiin liittyvillä tekijöillä. Organisaatiotason tarkasteluissa lähdetään siitä oletuksesta, että sairaala tai vanhainkoti on organisaatio, jonka toimintaperiaatteet ja toiminnassa näkyvät muutokset heijastavat organisaation johdon tavoittelemia päämääriä. Toinen lähestymistapa on ollut tarkastella suoraan päätöksentekijöitä organisaation sisällä, esimerkiksi lääkäreiden ja asiakkaiden päämies-agenttisuhdetta (Gaynor, 1994; T. McGuire, 2000), lääkäreiden luomaa kysyntää (supplier-induced demand) (Roemer & Shain, 1959) tai erilaisten rahoitus- ja kannustinjärjestelmien vaikutusta lääkäreiden käyttäytymiseen (Croxon, Propper, & Perkins, 2001).

On esitetty, että palveluntuottajien motiivit voittoa tavoittelemattomissa organisaatioissa johtaisivat tehottomuuteen (Newhouse, 1970; Worthington, 2004). Voittoa tavoittelemattomissa sairaaloissa päätöksentekoon voivat vaikuttaa esimerkiksi filantropia ja edistyksellisen teknologian suosiminen, jolloin sairaalat saattavat toimia tietoisesti tehottomasti. Voittoa tavoittelemattomat organisaatiot eivät myöskään välttämättä pysty minimoimaan kustannuksia, jos organisaatiolla ei ole korvaavia hyödykkeitä, joiden avulla tuloja voidaan lisätä. Julkisten palveluntuottajien toimintaan vaikuttavat myös toimintaympäristöön liittyvät rajoitteet ja lainsäädäntö. Yksittäisen vanhainkotosaston ja laitoksen näkökulmasta voidaankin ajatella, että ainakin osittain toimintamahdollisuudet, kuten esimerkiksi tulojen lisääminen lyhyellä aikavälillä ovat hyvin rajoitettuja. Esimerkiksi Blank ja Eggink (Blank & Eggink, 2001) olettivat tutkimuksessaan, että julkisen vanhustenhuollon laitoksen hoitopäivämäärä, budjetti ja asiakasrakenne ovat kaikki eksogeenisia, jolloin ainoa endogeeninen tekijä olisi palvelun laatu. Laatu olisi näin ollen keskeinen tekijä, jolla eri julkiset yksiköt kilpailevat palvelumarkkinoilla. Toisaalta edellä kuvatun tiedon epäsymmetrian ja laadun arvioimisen vaikeuden vuoksi julkisissa hoivapalveluissa kilpailulle ei välttämättä ole merkittäviä kannustimia. Laadun parantaminen ei myöskään ole välttämättä kannustavaa ilman siihen liittyvää yksilö- tai yksikkökohtaista palkitsemisjärjestelmää, jos lisälaadun tuottamisen koetaan lisäävän vain kustannuksia ja työmäärää. Tällöin laadun tuottaminen riippuu merkittävästi työntekijöiden ammattietiikasta.

Vanhustenhuollon talouteen liittyvissä tutkimuksissa palveluntuottajien käyttäytymismotiiveja ja toiminnan taloudellisia tavoitteita ei ole juurikaan arvioitu. Yrityksen teoriasta sairaalatoimintaan sovelletut tuotanto- ja talousteoreettiset mallit (esim. Linna, 1998) lienevät kuitenkin pääosin sovellettavissa myös vanhusten laitoshoidon. Empiiristen tutkimusten tai teorian perusteella ei kuitenkaan voida varmasti sanoa, mitkä ovat Suomessa pitkäaikaishoitoa tuottavien toimijoiden taloudellisen käyttäytymisen motiivit. Palvelun tuottajat voivat maksimoida tuotosmääränsä, maksimoida sekä tuotosta että hoidon laatua, minimoida kustannuksiaan tai Niskasen teorian mukaisesti pyrkiä maksimoimaan vuosittain budjettiaan. Päätöksenteko on käytännössä ilmeisesti monitavoitteista. Coellin (Coelli et al., 1998) mukaan julkisen organisaation toiminnan todennäköisimpänä tavoitteena voidaan pitää tuotosmäärän maksimointia annetun budjetin rajoissa, jolloin empiirisessä tuottavuustutkimuksessa saattaa olla tarkoituksenmukaisinta keskittyä pääasiassa teknisen tehokkuuden tutkimukseen. Teknisen tehokkuuden laskeamiseen ei tarvita hintainformaatiota eikä välttämättä edes käyttäytymisoletuksia.

4.6 Tuotannon määrän ja laadun vaihtosuhide

Toimintaympäristössä, jossa päätöksenteolla on useita tavoitteita, joudutaan tekemään valintoja erilaisten tavoitteiden saavuttamiseksi. Perinteisesti valinta voidaan tiivistää tehokkuuden ja oikeudenmukaisuuden tavoitteluun (Le Grand et al., 1992).⁵ Oikeudenmukaisuuden toteutuminen on yhteiskunnallinen tavoite ja siten ennen kaikkea yhteiskunnan ja poliittisen päätöksenteon tehtävä. Palvelutuotannossa tuottaja tavoittelee tehokkuutta optimoimalla tuotannon määrää ja laatua. Terveysthuollossa ajatellaan usein, että laadun lisääminen vaatii lisävoimavaroja ja että se maksaa, ja että tuottavuuden parantaminen johtaa väistämättä heikompaan laatuun. Tuotannon määrän ja laadun välistä suhdetta on havainnollistanut muun muassa Newhouse (Newhouse, 1970). Hän on esittänyt, että määrä ja laatu ovat tuotoksen ominaisuuksia, joiden tuottamiseen voittoavoitteluun päätöksentekijä voi kohdentaa voimavaroja. Voimavarojen niukkuudesta kuitenkin seuraa, että niiden välillä on vaihtosuhide (trade-off). Newhousen ajattelun mukaan tehokas yksikkö on sellainen, joka tuottaa sen hetkiselä teknologian tasolla sekä parasta mahdollista laatua että suurinta mahdollista määrää (maximand). Tehoton yksikkö voi siten tuottaa parempaa laatua tai enemmän määrää tai molempia yhtäaikaaisesti. Newhousen malli on sovellettavissa esimerkiksi vanhusten laitoshoidon. Newhousen mallia on havainnollistettu kuviossa 5 laadun ja tuotannon määrän vaihtosuhidekäyrän (trade-off curve) avulla. Tuottaja voi tuotantomahdollisuuksien rajoissa valita tuotetun laadun ja määrän kombinaation. Piste A



Lähde. Newhouse 1970

KUVIO 5. Laadun ja määrän vaihtosuhidekäyrä

⁵ Le Grand (Le Grand et al., 1992) kuten Barr (Barr, 1998) tarkoittavat tehokkuustavoitteella tässä yhteydessä niukkojen voimavarojen optimaalista käyttöä hyvinvoinnin, terveyden ja elämänlaadun (maximand) aikaansaamiseksi. Lisäksi Barr (emt.) liittää tehokkuusmääritelmäänsä myös ihmisten preferenssit.

kuvaa kohtaa, jossa lisälaadun tuottaminen on lisännyt kysyntää ja tuotettua määrää yhä vähemmän, kunnes lisälaadun tuottaminen on johtanut yhä korkeampiin kustannuksiin ja tuotetun määrän vähenemiseen. Tällöin vaihtosuhdekäyrä on taipunut taaksepäin. Pisteen A leikkaava indifferenssikäyrä (I) kuvaa pistettä, jossa tuottaja saavuttaa suurimman hyödyn. Valinta määrän ja laadun kombinaatioiden välillä seuraa tuottajan (yrityksen) maksimointiongelman ratkaisusta.

Vanhusten laitoshoidossa on luonnollisesti mahdollista, että parempi hoidon laatu voisi säästää ainakin asiakas- ja hoitojaksokohtaisia kustannuksia. Hoitovirheiden korjaaminen ja haitallisten hoitotulosten esiintyminen, kuten esimerkiksi painehaavat ja kaatumiset voivat pidentää asiakkaiden hoitojaksoja ja johtaa aikaisempaa koulutetumman henkilökunnan tarpeeseen ja kalliimpien hoitovälineiden ja -käytäntöjen käyttöön. Laadun ja kustannusten yhteyteen vaikuttaa kuitenkin määritelmällisesti se, mitä hoidon laadulla konkreettisesti ymmärretään. Eräiden laadun osa-alueiden, esimerkiksi toimintaympäristön viihtyisyyden parantaminen, nostaa todennäköisesti ainakin pääomakustannuksia, jolloin myös toiminnan kokonaiskustannukset nousevat ja tuottavuus heikkenee. Sitä vastoin asiakastytyväisyyttä voidaan mahdollisesti parantaa ilman lisäkustannuksia tai tuotannollisia tehokkuustappioita.

Yksikkö saattaa toimia Newhousen mallin mukaisesti tehottomasti, jos sillä ei ole sisäisiä tai ulkoisia kannusteita määrällisesti tehokkaaseen ja laadukkaaseen palvelutuotantoon. Perinteisesti kannustimilla tarkoitetaan työntekijöiden palkitsemiseen ja toiminnan rahoitukseen liittyviä menettelyjä. Satunnaiset häiriöt hoitopäivien tuotannossa, puutteet henkilöstön johtamisessa ja voimavarojen allokoinnissa ja tehottomuutta aiheuttavat valinnat (mielitymykset) saattavat kuitenkin aiheuttaa tehottomuutta. Toiminnan laatukin voi vaikuttaa tuotannolliseen tehokkuuteen, jos laatu koetaan tuotoksen ominaisuuden sijaan itse tuotantoprosessiin liittyväksi tekijäksi, joka vaikuttaa siihen, miten voimavaroja kytetään muuttamaan suoritteiksi. Tällöin laatua ei operationalisoida tuotoksen ominaisuudeksi, vaan voimavarojen ja prosessien ominaisuudeksi. Prosessien ja voimavarojen laadulliset ominaisuudet saattavat olla sellaisia, että ne vaikuttavat tuotannolliseen tehokkuuteen. Vanhusten laitoshoidossa esimerkiksi kuntouttavan työotteen soveltaminen ja erityyppiset interventiot saattavat edellyttää henkilöstömäärän lisäämistä ja työkäytäntöjen muuttamista, jolloin ei välttämättä pystytä saavuttamaan tuotantoteknologian mahdollistamaa teknisesti tehokkaan tuotannon rintamaa. Yksikkö voi myös tietoisesti panostaa tavanomaista enemmän henkilöstön määrään tai tietynlaiseen henkilöstön ammattirakenteeseen (voimavaralaatuun), vaikka lisäyksellä ei saavutettaisikaan vastaavaa lisäystä vaikuttavuuteen tai tuotettuihin hoitopäiviin.

5 Tuottavuuden ja tuotannollisen tehokkuuden mittaaminen

5.1 Terveydenhuollon tuottavuuden mittaaminen

Tuotantoteoriaan perustuvaa terveydenhuollon tuottavuuden mittaamista voidaan kritisoida siitä, ettei siinä pystytä riittävästi huomioimaan toiminnan hyvinvointivaikutuksia ja laadullisia tekijöitä. Tuottavuuden parantuminen voi olla seuraus teknologisen kehityksen lisäksi tuotannontekijöiden määrän ja laadun muutoksesta. Tuottavuuden muutosta ei tulisikaan koskaan suoraviivaisesti tulkita suhteessa hyvinvointivaikutuksiin, jotka ovat terveystalvelujen perimmäisenä tavoitteena. Viime aikoina tuottavuuden mittaamista varten on pyritty kehittämään menetelmiä, jossa toiminnan laadulliset muutokset voitaisiin ottaa huomioon. Esimerkiksi sairaaloiden tuotoksen muutosta voitaisiin painottaa laatu-painotteisilla elinvuosilla, jonotusajalla ja kuolleisuudella, eikä perinteiseen tapaan painottamalla tuotoksia niiden yksikkökustannuksilla (ks. Dawson et al., 2004). Tällöin joudutaan kuitenkin todennäköisesti valitsemaan suhteellisen pieni määrä laatu- tai vaikuttavuusmittareita. Lisäksi voidaan kysyä, onko tällöin kyse enää varsinaisesti tuottavuuden mittaamisesta vai pitäisikö puhua pikemminkin kustannus-vaikutavuuden mittaamisesta? Toisena esimerkkinä uusista innovaatiosta voidaan mainita Stakesin ja viiden sairaanhoitopiirin PERFECT (PERFORMANCE, Effectiveness and Cost of Treatment episodes) -hanke. Tämän hankkeen tavoitteena on luoda indikaattoreita ja malleja, joiden perusteella voidaan systemaattisesti seurata eri rekisteri- ja tilastotietoihin perustuen erikoissairaanhoiton hoitoketjujen vaikuttavuutta, laatua ja kustannus-vaikuttavuutta alueiden, sairaaloiden ja väestöryhmien kesken. Tavoitteena on myös tarkastella, mitkä tekijät selittävät alueellisia ja tuottajakohtaisia eroja koskien ensivaiheissa laadittuja laatuindikaattoreita. Hankkeessa käytetään vaikuttavuusmittarina muun muassa kuolleisuutta.

Tässä tutkimuksessa käytetään vain poikkileikkausaineistoa, jonka avulla voidaan spesifioida tehokkuusrintama, mutta ei voida arvioida teknologisen muutoksen merkitystä. Seuraavaksi käsitellään taloustieteellisissä tuottavuustutkimuksissa käytettyjä poikkileikkausaineistoon soveltuvia tuotannollisen tehokkuuden mittaamisen menetelmiä, jotka voidaan jakaa parametrisiin ja epäparametrisiin. Tämän jälkeen luvussa 6 esitetään, millä tavoin näissä menetelmissä voidaan ottaa huomioon toiminnan laadulliset ulottuvuudet.

5.2 Parametrinen lähestymistapa

Perinteisten ekonometristen tuotanto- tai kustannusfunktioiden estimoinnin ohella tuotannollisen tehokkuuden mittaamiseen on käytetty useimmiten panos-tai tuotosetäisyysfunktioihin perustuvia menetelmiä. Etäisyysfunktioihin (Farrell, 1957; Shepard, 1970) perustuvien menetelmien avulla voidaan identifoida tehokkuusrintama ja arvioida eri yksiköiden tehokkuutta suhteessa tähän rintamaan. Tulokseksi saadaan kullekin yksikölle tehokkuusluku, joka kertoo etäisyyden tehokkuusrintamaan. Tehokkuusrintama määrittyy tehokkaiden eli ei-dominoitujen yksiköiden perusteella ja tehottomat yksiköt jäävät tämän rintaman alapuolelle. Kuten jo aiemmin kävi ilmi, tuotantoteoriassa ja sen menetelmissä tehokkuus (efficiency) on nimenomaan etäisyysfunktioihin liittyvä käsite. Tästä syystä tässä luvussa 5 ja kappaleessa 6.1 termeillä tehokkuus, tehottomuus ja tehokkuusluku viitataan yksikön etäisyyteen tehokkuusrintamasta.

Parametrinen lähestymistapa perustuu tuotanto- ja kustannusfunktioiden estimointiin ja ekonometrian tutkimusperinteeseen (Shepard, 1970). Tällöin oletetaan, että tuotantofunktio on tunnettu, jolloin sen parametrit voidaan estimoida tilastollisesti. Esimerkiksi tavallista pienimmän neliösumman regressioanalyysia (PNS) voidaan käyttää tuotantofunktion estimointiin. Tällöin voidaan kuitenkin arvioida vain panosten ja tuotosten keskimääräistä riippuvuutta (Coelli et al., 1998; Cooper, Seiford, & Tone, 2000). PNS-menetelmällä ei voida estimoida tehokkuusrintamaa eikä erottaa tehottomuutta satunnaisvaihtelusta. PNS:n avulla estimoitua mallia voidaan korjata muuttamalla vakiota siten, että virhetermit ovat positiivisia, jolloin puhutaan COLS-menetelmästä (Corrected Ordinary Least-Squares regression). Tämä menetelmää ei kuitenkaan yleisesti käytetä, koska on olemassa muita menetelmiä, joiden avulla sekä parametristimaatit, satunnaisuus ja tehottomuus pystytään määrittelemään yhtäaikaaisesti.

Deterministinen rintamamalli on etäisyysfunktioihin perustuva sovellus, jonka avulla voidaan arvioida yksiköiden etäisyyttä tehokkuusrintamassa oleviin tehokkaihin yksiköihin (Coelli et al., 1998). Tällöin kuitenkin joudutaan olettamaan, että havaittu ero tehokkuusrintamaan on kokonaan tehottomuutta, jolloin myös joudutaan olettamaan, ettei aineistossa ole satunnaisvaihtelua. Satunnaisvaihtelun puuttuminen ei kuitenkaan tavallisesti ole perusteltu oletus ekonometrisissa mal-leissa. Tähän ongelmaan on kehitetty stokastinen rintamamalli, joka on myös etäisyysfunktioihin perustuva sovellus. Sen avulla voidaan estimoida tehottomien yksiköiden etäisyys tehokkuusrintamasta kun sekä satunnaisuus että tehottomuus ovat huomioitu (Aigner, Lovell, & Schmidt, 1977; Meeusen & van den Broeck, 1977). Stokastisia rintamamalleja voidaan käyttää tuotanto- ja kustannusfunktioiden estimoimiseen sekä poikkileikkaus- että paneeliaineistoista. Suomessa tätä menetelmää on sovellettu muun muassa vanhusten palveluissa (Björkgren, Häkkinen, &

Linna, 2000) ja sairaaloiden (Linna & Häkkinen, 1999) tuotannolliseen tehokkuuteen liittyvissä tutkimuksissa.

Stokastisen rintamamallin keskeisin ero tavalliseen regressiomalliin ja deterministiseen lähestymistapaan on siinä, että virhetermi on jaettu kahteen komponenttiin: tavalliseen satunnaiseen virhetermiin ja tehottomuustermiin (Greene, 1993). Ne voidaan myös estimoida yhtäaikaaisesti. Yleinen lineaarinen esitys stokastiselle tuotantofunktiolle on

$$y_i = \alpha + \sum_j \beta_j \cdot x_j + v_i - u_i, \quad (1)$$

jossa y_i on tuotos, β_j on estimoitu parametri tuotannontekijälle x_j , v_i on satunnainen virhetermi ja u_i on positiivinen tehottomuustermi. Virhetermi v on normaalisti jakautunut satunnaismuuttuja. Tehottomuustermi u on positiivinen satunnaismuuttuja eli

$$\begin{aligned} v_i &\sim N[0, \sigma_v^2] \\ u_i &= |U|, U \sim N[0, \sigma_u^2] \end{aligned} \quad (2)$$

Mallin virhetermin jakautuminen kahteen osatekijään riippuu u :n oletetusta jakaumasta. Sen katsotaan tavallisimmin noudattavan puolinormaalaa jakaumaa, katkaistua normaalijakaumaa, eksponentiaalista jakaumaa tai gammajakaumaa (Coelli et al., 1998). U :n varianssin poikkeama nolasta ($\sigma_u^2 \neq 0$) on merkki tehottomuudesta. Tehottomuuden osuus saadaan virhetermien hajontojen suhteesta.

Stokastisia rintamamalleja voidaan käyttää myös kustannusfunktioiden estimointiin. Stokastinen kustannusfunktio on muotoa

$$C_i = C(y, w) + v_i + u_i, \quad (3)$$

jossa C_i on kokonaiskustannukset, y on tuotos, w on panoshinta, v_i on satunnainen virhetermi ja u_i on positiivinen tehottomuustermi. v ja u ovat jakautuneet kuten yllä.

Stokastiset rintamamallit edellyttävät yleistetyn lineaarisen mallin tapaan mallin funktiomuodon spesifiointia. Tämän lähestymistavan heikkoudet ovatkin juuri mallin funktionaalisen muodon ja tehottomuuteen liittyvän virhetermin määrittämisen vaikeus. Tehottomuustermiin jakaumaa ei voida käytännössä juurikaan testata. Jakauman valinta saattaa eräissä tapauksissa vaikuttaa yksiköiden tehokkuuslukuihin. Yksi ratkaisu jakaumaoletuksiin liittyviin ongelmiin on soveltaa bayesilaisia malleja, joita on viime vuosina käytetty myös terveystaloustieteellisessä tutkimuksessa (O'Hagan & Luce, 2003). Stokastinen rintamamalli estimoidaan käyttäen suurimman uskottavuuden menetelmää (maximum likelihood), jossa

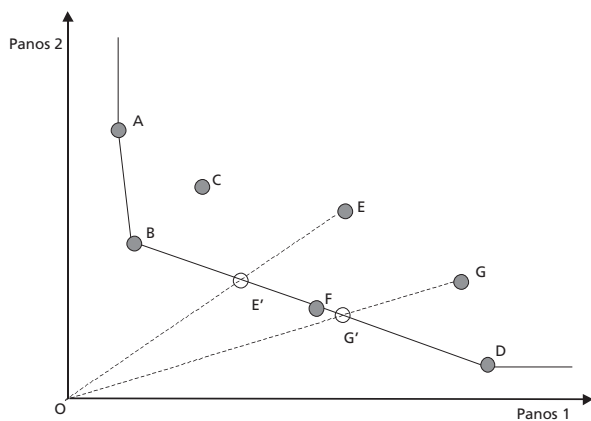
epälineaarisen ohjelmoinnin avulla haetaan parametreille aineistoon parhaiten sopivimmat (todennäköisimmät) estimaatit.

5.3 Epäparametrinen lähestymistapa

Vaihtoehtoinen etäisyysfunktioihin perustuva lähestymistapa on Data Envelopment -analyysi (DEA), jonka kehitti Farrell 1950-luvulla (Farrell, 1957). Charnes, Rhodes ja Cooper (Charnes, Cooper, & Rhodes, 1978) kehittivät sitä laskennallisesti empiiriseen tutkimukseen paremmin soveltuvaksi, jolloin menetelmän käyttö myös yleistyi. DEA:ssa saman toiminnallisen tavoitteen omaavien yksiköiden havaintojoukosta karakterisoidaan tehokkuusrintama, jossa lähtökohtana ovat todellinen voimavarojenkäyttö (panokset) ja todelliset tuotokset. Menetelmä ei perustu ekonometriseen estimointiin eikä siihen, että tuotantofunktio olisi tunnettu, vaan aineistolähtöiseen lineaarisen ohjelmointiin. DEA ei siten edellytä stokastisissa rintamamalleissa tehtäviä oletuksia mallin funktionaalisesta muodosta tai tehottomuuserin jakaumasta. DEA-malli voidaan laskea joko tuotos- tai panosorientoituneesti. Menetelmää on sovellettu laajasti Suomessa sairaaloiden, terveyskeskusten ja vanhusten pitkäaikaishoitolaitosten tuottavuustutkimuksissa (Björkgren, Häkkinen, & Linna, 2001; Linna & Häkkinen, 1999; Luoma, Järviö, Suoniemi, & Hjerppe, 1996).

Kuviossa 6 on havainnollistettu graafisesti yksiköiden teknistä tehokkuutta ja tehokkaiden yksiköiden muodostamaa tehokkuusrintamaa. Yksiköt tuottavat saman tuotoksen käyttämällä tuotannossa eri tavoin kahta eri panosta. Konvekssit kombinaatiot (lineaarikombinaatiot) salliva tehokkuusrintama muodostuu origoa lähimpänä olevista havainnoista (A, B, F ja D), joiden DEA:n avulla laskettu tehokkuusluku on yksi. Toisin sanoen nämä yksiköt ovat ei-dominoituja. Yksiköt tuottavat saman tuotoksen eri panosyhdistelmillä. Sen sijaan havainnot C, E ja G ovat teknisesti tehottomia, koska niiden pitäisi pystyä tuottamaan enemmän tuotosta käyttämällä panosmäärällä. Esimerkiksi yksiköiden E ja G tehottomuus tarkoittaa havaintojen etäisyyttä tehokkuusrintamasta eli etäisyyttä E,E' ja G,G'. Havainnon E tehokkuusluku saadaan siten suhteesta OE'/OE. On huomattava, että tehottomien yksiköiden referenssiyksiköt ovat tehokkaat virtuaaliyksiköt (esim. G'), jotka sijoittuvat identifioituun tehokkuusrintamaan.

DEA-menetelmää voidaan käyttää stokastisten rintamamallien tavoin teknisen tehokkuuden ja kustannustehokkuuden laskemiseen. Käsiteltäessä paneelidataneistoa käytetään Malmqvist-indeksiä, joka on DEA:n sovellus. Yleisesti ottaen tutkimuksissa on havaittu, että DEA ja stokastiset rintamamallit tuottavat suhteellisen samansuuntaisia tehokkuuslukuja (Björkgren et al., 2000; Jacobs, 2000; Linna & Häkkinen, 1999).



KUVIO 6. Tekninen tehokkuus ja tehokkuusrintama

6 Tuottavuuserojen tutkimus

6.1 Tuotannollisen tehokkuuden erojen selittäminen

Tuottavuustutkimuksissa havaitaan usein suuriakin eroja yksiköiden välillä. Suomessa vanhustenhuollon tuotannolliseksi tehottomuudeksi on arvioitu aiemmin noin 20 prosenttia (Björkgren et al., 2000; Laine, 2003), sairaalatoiminnan tuotannolliseksi tehottomuudeksi noin 13 prosenttia (Linna & Häkkinen, 1999) ja terveyskeskusten tuotannolliseksi tehottomuudeksi noin 20 prosenttia (Aaltonen, Järviö, Luoma, & Rätty, 2004). Yleensä empiirisissä tutkimuksissa ollaan kuitenkin kiinnostuneita tuotannollisen tehottomuuden syistä eikä vain yksiköiden välisistä eroista tai toimialan tuottavuudesta. Tehokkuuslukujen laskemisen menetelmät ovat nykyisin varsin vakiintuneita (ks. edellinen luku). Tuotannollisen tehottomuuden syiden ja tehokkuuserojen syvälliseen arviointiin käytettävät menetelmät ja lähestymistavat ovat kuitenkin itse tehokkuuslukujen laskentaa vakiintumattomampia.

Selitetessä yksiköiden välisten tehokkuuserojen syitä on käytetty joko yksivaiheista tai kaksivaiheista lähestymistapaa. Kaksivaiheisen lähestymistavan ensimmäisessä vaiheessa yksiköille lasketaan tehokkuusluvut joko DEA:n tai stokastisen rintamamallin avulla. Toisessa vaiheessa tehokkuusluvuissa esiintyvää varianssia selitetään joko PNS:n tai Tobit-regression avulla. Tobit-regression käyttöä on suositeltavaa PNS:n sijaan, koska esimerkiksi teknisen tehokkuuden tehokkuuslukujen jakauma on katkaistu välillä 0–1 (1 on tehokas ja < 1 on tehoton). Kaksivaiheisen analyysin toisessa vaiheessa joudutaan olettamaan, että tehottomuusvirhetermit (tehokkuusluvut) ovat riippumattomasti ja yhtäläisesti jakautuneita. Tämä ei kuitenkaan päde, jos niiden oletetaan samanaikaisesti olevan funktio tietyistä yksiköitä kuvaavista muuttujista. Tällöin toisen vaiheen parametriesiimaatit ovat harhaisia (Coelli et al., 1998). Kumbhakar ym. (Kumbhakar, Ghosh, & McGuckin, 1991) ja Reifschneider ym. (Reifschneider & Stevenson, 1991) ovat tämän harhaisuuden välttämiseksi spesifioineet yksivaiheisen lähestymistavan. Siinä tehottomuusluvut ja tehottomuuseffektit (tehottomuuden syyt) voidaan estimoida yhtäaikaaisesti stokastisella rintamallilla käyttäen suurimman uskottavuuden menetelmää. Battese ja Coelli (Battese & Coelli, 1993, 1995) laajensivat mallia paneeliaineistoon soveltuvaan. Vallitsevan käsityksen ja simulaatiotutkimusten perusteella yksivaiheista estimointia pidetään selvästi kaksivaiheista suositeltavampana (Simar & Wilson, 2003; Wang & Schmidt, 2002). Viimeaikaisissa parametrisia menetelmiä soveltavissa tutkimuksissa ei enää ole juurikaan käytetty kaksivaiheista lähestymistapaa. Yksivai-

heinen stokastinen rintamamalli esimerkiksi teknisen tehokkuuden estimoimiseksi voidaan esittää seuraavasti

$$y_i = \alpha + \sum_j \beta_j \cdot x_j + v_i - u_i, \quad (4)$$

jossa y_j, β_j, x_j, v_i ovat yhtäläisiä tavallisen stokastisen rintamamallin kanssa. Tehottomuuseffektit u_i määritellään kuitenkin seuraavasti:

$$u_i = \delta Z_i + w_i, \quad (5)$$

jossa Z_i on tehottomuuteen liittyvien muuttujien vektori, δ on vastaavien muuttujien parametreille estimoitava vektori ja w_i on riippumattomasti jakautunut virhetermi. Erona tavalliseen stokastiseen rintamamalliin (1) se, että yksivaiheisessa mallissa oletetaan, että tehottomuus u_i on funktio erikseen spesifioiduista muuttujista.

Yllä esitettyä mallispesifikaatiota (4) on käytetty useissa viimeaikaisissa empiirisissä tutkimuksissa, joissa tarkoituksena on ollut tutkia tehottomuutta aiheuttavia tekijöitä (Li & Rosenman, 2001; Linna, 1998; Rosko, 2001, 2004). Tähän läheisesti liittyvä lähestymistapa on ollut se, että tehottomuuteen yhteydessä olevat tekijät sijoitetaan suoraan kustannusfunktioon (tai tuotantofunktioon). Tällöin malli voidaan estimoida perinteisen kustannusfunktioimallin avulla, jolloin voidaan arvioida, miten tekijät ovat yhteydessä kustannuksiin (esim. Crivelli, Filippini, & Lunati, 2002; Folland & Hoffer, 2001; Knox, Blankmeyer, & Stutzman, 2003; Zuckerman, Hadley, & Iezzoni, 1994). Se, päätetäänkö tehottomuuseffektit sisällyttää suoraan kustannusfunktioon vai erillisiksi tekijöiksi tehottomuuskomponenttiin, on ennen kaikkea tutkimusasetelmaan ja käytettävissä oleviin muuttujiin liittyvä valinta. Tällöin on selvítettävä, halutaanko tarkastella, mitkä tekijät vaikuttavat tehokkuusrintaman määräytymiseen vai halutaanko tehottomuuseffekteillä havainnollistaa tekijöitä, jotka vaikuttavat yksiköiden etäisyyteen tehokkuusrintamasta.

DEA-menetelmän avulla laskettujen tehokkuuslukujen selittämiseen ei ole vielä vakiintunutta menettelyä kaksivaiheisuuden tilalle. Yksivaiheista estimointia ei voida soveltaa epäparametrisessa analyysissä. Tästä syystä useissa tutkimuksissa on edelleen sovellettu kaksivaiheista lähestymistapaa, jos tehokkuusluvut on laskettu DEA-menetelmää käyttäen (esim. Björkgren et al., 2000; Linna, Nordblad, & Koivu, 2003). DEA-menetelmää soveltavissa tutkimuksissa yksiköiden välisten tehokkuuserojen syitä onkin usein tarkasteltu yksinkertaisilla kuvailevilla menetelmillä, esimerkiksi vertaamalla maantieteellisten alueiden tai erikokoisten yksiköiden tehokkuuslukujen keskiarvoa ja hajontaa (esim. Aaltonen et al., 2004).

6.2 Vanhustenhuollon tuotannollisen tehokkuuden ja laadun empiiriset tutkimukset

Vanhustenhuollon tuotannolliseen tehokkuuteen, kustannuksiin ja laatuun liittyvien tutkimusten määrä on huomattavasti niukempi kuin esimerkiksi sairaaloiden toimintaa käsitelleiden tutkimusten. Ensimmäiset vanhusten pitkäaikaishoidon tuotannolliseen tehokkuuteen liittyvät tutkimukset tehtiin Yhdysvalloissa 1980-luvulla. Tuolloin keskeisenä tutkimuskohteena olivat sekä julkisten ja yksityisten laitosten välisten tehokkuuserojen arvioiminen epäparametrisilla menetelmillä että rahoitusjärjestelmän toimivuuden tutkiminen kohonneiden terveydenhuoltomenojen johdosta. Sexton ym. (Sexton, Leiken, Sleeper, & Coburn, 1989) havaitsivat 1980-luvulla DEA-menetelmää käyttäen, että prospektiiviseen rahoitusjärjestelmään siirtyminen lisäsi pitkäaikaishoitolaitosten henkilöstöresursseja, jolloin laitosten tuotannollinen tehokkuus heikkeni. Sextonin ym. mukaan syitä tuotannollisen tehokkuuden heikkenemiseen saattoivat olla myös voimavaralaadun heikentyminen, laadukkaampien tuotosten aikaansaaminen tai aiempaa huonompi johtaminen. Nyman ja Bricker (Nyman & Bricker, 1989) puolestaan arvioivat, että Yhdysvalloissa 1980-luvulla voittoa tavoittelevat vanhustenhoitolaitokset olivat tuotannollisesti tehokkaampia kuin voittoa tavoittelemattomat yksiköt.

Viime vuosikymmenen aikana vanhustenhuollon tuotannolliseen tehokkuuteen liittyvien tutkimusten määrä lisääntyi erityisesti Yhdysvalloissa ja samalla tutkimuksia laajennettiin menetelmällisesti soveltamalla myös parametrisia menetelmiä (Vitaliano & Toren, 1994). Julkisten ja yksityisten laitosten tuotannollisen tehokkuuden vertailu oli edelleen 1990-luvulla tärkeimpiä tutkimuskysymyksiä. Tutkimusten tulokset ovat pääosin tukeneet sitä, että ainakin Yhdysvalloissa yksityiset vanhustenhuoltolaitokset olisivat julkisia laitoksia tehokkaampia (Chattopadhyay & Ray, 1996; Ozcan, Wogen, & Mau, 1998; Rosko, Chilingerian, Zinn, & Aaronson, 1995). Omistusmuodon vaikutusten lisäksi on tutkittu myös ketjuuntumisen ja kilpailuolosuhteiden yhteyttä tuotannolliseen tehokkuuteen. Eräissä tutkimuksissa on havaittu, että kilpailun lisääminen ja tuottajien vähäinen keskittyminen lisäävät tuotannollista tehokkuutta (Anderson, Lewis, & Webb, 1999; Blank & Eggink, 2001). Mitchellin ym. (Mitchell, Venkatraman, Banaszak-Holl, Baum, & Berta, 2003) mukaan Yhdysvalloissa ketjuuntuneissa vanhustenhoitolaitoksissa on hieman pienempi henkilöstömitoitutus ja hieman enemmän hoidon laatuun liittyviä puutteita. Voittoa tavoittelemattomissa laitoksissa oli puolestaan parempi laatu mutta myös suuremmat kustannukset kuin voittoa tavoittelevissa laitoksissa. Mitchell ym. korostavat, että vanhustenhuollon markkinat ovat Yhdysvalloissa kuitenkin segmentoituneet muun muassa asiakasrakenteen mukaan, mikä käytännössä rajoittaa laitoksen siirtymistä segmentistä toiseen. Segmentoituminen saattaa myös merkitä sitä, että asiakasvalikoituvuus asettaa tiettyntyyppiset laitokset epäedulli-

seen asemaan tuottavuus- ja laatuvertailuissa. Suomessa julkisten ja yksityisten laitosten vertaileminen ei ole toistaiseksi yhtä keskeinen kysymys kuin Yhdysvalloissa, sillä Suomessa lähes koko vanhainkotihoito ja terveyskeskusten vuodeosastohoito ovat julkisen sektorin vastuulla (SVT, 2003). Yksityiset palvelut ovat laajentuneet toistaiseksi erityisesti kotipalveluissa ja palveluasumisessa (Kauppinen, 2001). Suomessa vanhustenhuollon markkinat ovat jossain määrin segmentoituneet. RAI-hankkeessa saatujen viimeaikaisten tulosten mukaan julkisissa vanhainkodeissa ja palvelutaloissa olevien asiakkaiden hoito vaatii keskimäärin enemmän voimavaroja kuin yksityisissä laitoksissa olevien asiakkaiden hoito. Tulevaisuudessa Suomessa on kuitenkin varmasti tarkoituksenmukaista ja tarpeellista arvioida sekä julkisen että yksityisten yksiköiden tuottavuutta ja hoidon laatua.

Omistusmuodon ja kilpailuolosuhteiden lisäksi tuotannollisen tehokkuuden eroja on pyritty selittämään yksikön koolla ja muilla tekijöillä, jotka kuvaavat organisaatiota. Tutkimukset eivät kuitenkaan yksiselitteisesti tue sitä, että suuret laitokset olisivat pieniä tehokkaampia (Björkgren et al., 2001; Kooreman, 1994; Laine, Linna, Häkkinen, & Noro, 2005; Sexton et al., 1989; Vitaliano & Toren, 1994). Eräät tutkimukset ovat päättyneet siihen, että osastolle tai laitokselle olisi kuitenkin olemassa jokin optimikoko. Esimerkiksi Björkgrenin ym. (Björkgren et al., 2001) mukaan tuotannon mittakaavaedun kannalta optimaalinen vuodeosastokoko olisi Suomessa noin 35 hoitopaikkaa. Yksikön koon lisäksi edellä mainituissa tutkimuksissa korkeaa tuotannollista tehokkuutta selittäviksi tekijöiksi on havaittu muun muassa korkea kuormitusaste, laitostyyppi, yksikön maantieteellinen sijainti ja henkilöstön ammattirakenne (perus- ja lähihoitajien korkea osuus henkilöstöstä).

Olellainen jatkokysymys yllä kuvattujen tulosten perusteella onkin se, aiheutuvatko yksiköiden väliset tuottavuuserot toimintaympäristöön ja omistusmuotoon liittyvistä tekijöistä vai hoidossa olevien asiakkaiden erilaisesta voimavaratarpeesta? Runsaasti voimavaroja vaativan asiakasrakenteen on osoitettu useassa tutkimuksessa olevan yhteydessä korkeisiin kustannuksiin (Blank & Eggink, 2001; Knox et al., 2003; Ozcan et al., 1998). Tästä syystä asiakasrakennetta on tuottavuustutkimuksissa käytetty myös tuotosta (hoitopäiviä) vakioivana tekijä eikä niinkään aina yksiköiden välisiä tuotannollisen tehokkuuden eroja selittävänä muuttujana (Björkgren et al., 2000). Tällä tavoin eri yksiköiden tuotokset on saatu yhteismittaliseksi, joka on edellytyksenä luotettavalle tuottavuustutkimukselle (Newhouse, 1994).

Vanhustenhuollossa tuotannollisen tehokkuuden, kustannusten ja laadun yhteys on olellainen osa toiminnan arviointia. Tuotannolliseen tehokkuuteen liittyvät vertailut ja tutkimukset ovat ongelmallisia, jos toiminnan laatua ei ole mitenkään huomioitu (Newhouse, 1994). Laatu on ollut jossakin määrin mukana eräissä vanhustenhuollon tuottavuustutkimuksissa, mutta tulokset ovat olleet ristiriitaisia. Vitaliano ja Toren (Vitaliano & Toren, 1994) arvioivat, että laatu mitattuna henkilöstömäärällä ei ollut yhteydessä tekniseen tehokkuuteen Winconsinin van-

hainkodeissa Yhdysvalloissa. Rosko (Rosko et al., 1995) päätyi samaan johtopäätökseen käyttäessään laadun mittarina voimavaralaadun sijaan joitakin prosessi- ja tuloslaatuindikaattoreita. Tutkimuksen mukaan Yhdysvalloissa vanhainkotien tehokkuuserot selittyivät laadun sijaan johtamiseen ja toimintaympäristöön liittyvillä tekijöillä, kuten esimerkiksi kuormitusasteella, rahoituslähteellä ja omistumuodolla. Tulos on yhdenmukainen Knoxin ym. (Knox et al., 2003) havaintojen kanssa. Heidän mukaansa laadun vaikutukset yksikön kokonaiskustannuksiin ovat hyvin vähäiset verrattuna hallintoon ja toimintaympäristöön liittyvien tekijöiden vaikutuksiin. Knoxin ym. tutkimuksessa laadun ja liikevoiton havaittiin jopa paranevan samanaikaisesti. Weech-Maldonadon ym. (Weech-Maldonado, Neff, & Mor, 2003) mukaan vanhainkodit, jotka tuottivat laadukkaampaa hoitoa, saavuttivat alhaisemmat asiakaskohtaiset kustannukset ja paremman taloudellisen kokonaistuloksen. Gertlerin ja Waldmanin (Gertler & Waldman, 1992) mukaan toimet laadun parantamiseksi voivat kuitenkin olla kalliita, mutta kokonaiskustannusten vähentäminen on mahdollista ilman suuria muutoksia laadussa. Blank ja Eggink (Blank & Eggink, 2001) sovelsivat Gertlerin ja Waldmanin (Gertler & Waldman, 1992) Yhdysvalloissa kehittämää kustannusfunktiomallia julkisesti tuotetun vanhusten laitoshoidon arviointiin Hollannissa. Heidän mukaansa hoidon laatu on negatiivisesti yhteydessä panoshintoihin eli työvoiman hinnan kallistuessa hoidon laatu heikkenee. Tutkimustulos ilmeisesti kertoo siitä, että kohonneita henkilöstökustannuksia voidaan kompensoida heikentämällä laatua. Chou (Chou, 2002) puolestaan havaitsi, että tiedon epäsymmetrian takia voittoa tavoittelevissa pitkäaikaishoitolaitoksissa hoidon laatu saattaa heikentyä kustannusten minimoinnin takia. Tällaista ongelmia ei välttämättä ole voittoa tavoittelemattomissa yksiköissä omistumuodon ja käyttäytymismotiivien erilaisen luonteen vuoksi.

Laadun on havaittu olevan tuotannollista tehokkuutta selvemmin yhteydessä eri henkilöstöryhmien määrään ja ammattirakenteeseen. Dellefield (Dellefield, 2000) on todennut laajan kirjallisuuskatsauksen perusteella, että suuri hoitajien määrä vanhainkodeissa olisi yhteydessä parempaan hoidon laatuun, mutta henkilöstön ammattirakenteen yhteys laatuun on epäselvä. Harrington ym. (Harrington et al., 2000) puolestaan havaitsivat, että mitä vähemmän vanhusten pitkäaikaishoitolaitoksessa oli sairaanhoitajien, hoitoapulaisten ja hallinnon henkilötöyvuosia, sitä enemmän esiintyi puutteita hoidon laadussa. Laitoksen toimintaympäristö ja muut alueelliset tekijät selittävät kuitenkin henkilötöyvuosia voimakkaammin näitä puutteita. Samansuuntaiseen tulokseen päätyi Yhdysvalloissa myös The General Accounting Office (GAO), jonka toimittamassa raportissa todetaan, että tutkimustulokset vanhusten laitoshoidon kustannusten ja laadun yhteydestä ovat ristiriitaisia. Mitä enemmän asiakkaat saivat hoitoaikaa, sitä parempi oli hoidon laatu (GAO, 2002). Hoidon laatu olisi siten yhteydessä siihen, kuinka hyvin henkilöstö pystyy kohdentamaan käytettävissä olevan työaikansa asiakkaille ja välittömään hoitotyöhön.

Yksiselitteisten johtopäätösten tekeminen aiemmista tutkimuksista on vaikeaa tulosten moninaisuuden takia. Erisuuntaisia tuloksia selittävät ilmeisesti muun muassa maaspesifit aineistot, erilaiset menetelmät ja tutkimusasetelmat ja erilaiset laatu- ja asiakasrakennemittarit. Laatu on myös moniulotteinen ilmiö, jonka sisällyttäminen täysin kattavasti empiiriseen tutkimukseen on vaikeaa. Aiemmat tutkimukset osoittavat kuitenkin sen, että tuotannollisen tehokkuuden eroja selitettäessä tulee ottaa huomioon niin toiminnan organisointiin, toimintaympäristöön kuin laatuunkin liittyvät tekijät.

7 Tutkimuksen tavoitteet

Vanhusten laitoshoidon tuotannollisen tehokkuuden ja laadun välistä yhteyttä on tutkittu kansainvälisesti suhteellisen vähän. Suomessakin on aiemmin tehty vain muutamia tutkimuksia vanhusten laitoshoidon tuotannollisesta tehokkuudesta (Björkgren et al., 2000, 2001; Lahtinen, 2003), mutta niissä ei ole huomioitu lainkaan hoidon laatuun liittyviä tekijöitä. Hieman kärjistetysti voidaan sanoa, että sosiaalitieteissä on toistaiseksi keskitytty pääasiassa vain hoidon laatuun liittyviin kysymyksiin huomioimatta kustannuksia ja tuottavuustutkimuksissa on vuorostaan keskitytty vain tuotannolliseen tehokkuuteen. Hoidon laatu sinänsä on luonnollisesti tärkeää, mutta se voi myös olla yhteydessä yksiköiden välisiin tuotannollisen tehokkuuden eroihin ja vaikuttaa yksiköiden suhteelliseen asemaan tuottavuusvertailuissa. Tuotannolliseen tehokkuuteen liittyvissä tutkimuksissa on otettava huomioon toiminnan laatuun liittyviä tekijöitä ja laatuun liittyvissä tutkimuksissa tulisi puolestaan arvioida myös kustannuksia. Tämä on erityisen tärkeää nykyisessä yhteiskunnallisessa tilanteessa, jossa sosiaali- ja terveyspalvelujen tuottavuuden parantaminen koetaan tärkeäksi. Samalla palvelujen tulisi olla kaikin tavoin laadukkaita. Päätöksenteon ja arvioinnin tueksi tarvitaan lisätietoa siitä, miten tuotannollinen tehokkuus ja hoidon laatu suhteutuvat toisiinsa vanhustenhoidossa ja millaisia eroja yksiköiden välillä on henkilöstövoimavarojen käytössä. Tämän kysymyksen tutkimiseen on Suomessa kansainvälisesti katsottuna nykyisin hyvät edellytykset hyvien tilastoaineistojen, rekisterien ja potilastietojärjestelmien vuoksi.

Tämän tutkimuksen yleisenä tavoitteena on tarkastella laadun ja tuotannollisen tehokkuuden välistä yhteyttä. Tutkimuksessa esitetään esimerkki siitä, miten tätä kysymystä voidaan sosiaali- ja terveydenhuollossa käsitteellistää ja tutkia. Empiirisenä tutkimuskohteena on vanhusten pitkäaikainen laitoshoido, jossa kysymys hoidon laadun ja tuotannollisen tehokkuuden välisestä yhteydestä on toimintaympäristön erityispiirteiden vuoksi erittäin tärkeä. Laitoshoidon erityispiirteiden (ks. luku 3), teorian (ks. luvut 4.5 ja 4.6) ja aiempien tutkimusten (ks. luku 6.2) perusteella tutkimuksen hypoteesina on oletus tuotannollisen tehokkuuden ja laadun välisestä vaihtosuhteesta. Tutkimuksessa odotetaan siten havaittavan yhteys korkean tuotannollisen tehokkuuden ja huonon hoidon laadun välillä. Henkilöstövoimavarojen käytön osalta voidaan odottaa, että asiakkaan saama hoitoaika (min/vrk) on erityisen suuri hoitopaikkamäärältään pienissä ja kodinomaisissa osastoissa.

Tutkimuksessa käytetyt empiiriset menetelmät ja tutkimusasetelma ovat mahdollisesti sovellettavissa erityyppisiin hoivapalveluihin. Tutkimustulokset on julkaistu neljässä kansainvälisessä terveystaloustieteen ja terveydenhuollon aikakausjulkaisussa julkaistussa artikkelissa. Tutkimuksen tavoitteet ovat seuraavat:

- 1) Arvioida pitkäaikaishoitoa tuottavien osastojen tuotannollista tehokkuutta ja sen vaihtelua (Osajulkaisu 1, 2 ja 3).
- 2) Arvioida, ovatko hoidon laadultaan heikoimmat osastot teknisesti tehokkaimpia (Osajulkaisu 1).
- 3) Arvioida, miten ja kuinka voimakkaasti hoidon laatu ja yksikön toimintaympäristö vaikuttavat tekniseen tehokkuuteen ja kustannustehokkuuteen (Osajulkaisu 2 ja 3).
- 4) Arvioida henkilöstön työajan käyttöön ja kohdentumiseen vaikuttavia osasto- ja yksilötasontekijöitä (Osajulkaisu 4).

8 Tutkimusaineistot

8.1 Tietolähteet

Tutkimuksessa käytettävät tiedot perustuvat Stakesin ja Chydenius-instituutin koordinoimaan RAI-tietojärjestelmän käyttöönotto ja pitkäaikaishoidon benchmarking -hankkeeseen, joka käynnistyi vuonna 2000. Tuolloin hankkeessa oli mukana noin 70 terveyskeskusta, vanhainkotiä ja palvelutaloä (Noro et al., 2001). Yhteistyökumppaneina hankkeessa toimivat myös Vanhustyön keskusliitto, Folkhälsan, Helsingin Diakonissalaitos ja Ikäinstituutti. Vuonna 2004 RAI-hankkeessa oli mukana jo noin 280 pitkäaikaisosastoa 25 kunnasta, mikä kattaa noin 20 prosenttia kaikista pitkäaikaishoidossa olevista asiakkaita. Asiakkaita näissä osastoissa oli vuonna 2004 noin 7 000 (Noro, 2005). Hankkeessa otettiin käyttöön RAI-tietojärjestelmä pitkäaikaishoidon asiakkaiden yksilöllistä arviointia ja hoidon laadun seurantaa varten. RAI on laaja kansainvälinen vanhustenhuollon arviointi- ja seurantajärjestelmä, joka kattaa asiakkaan koko hoito- ja palveluketjun aina akuuttihoitosta laitoshoitoon. RAI-järjestelmissä on yhteisiä kysymyksiä, joita voidaan hyödyntää varmistettaessa hoidon jatkuvuus hoitoketjun eri vaiheissa. RAI:n pätevyys ja luotettavuus on testattu useissa tutkimuksissa ja maissa (Björkgren, Häkkinen, Finne-Soveri, & Fries, 1999; Carrillo, Garcia-Altes, Peiro, & Portella, 1996; Hawes et al., 1995; Ikegami, Fries, Takagi, Ikeda, & Ibe, 1994). Suomessa pitkäaikaisen laitoshoidon ja kotihoidon RAI-järjestelmä on atk-pohjainen ja se on otettu käyttöön osana potilastietojärjestelmää. RAI-järjestelmä muodostuu MDS-lomakkeesta (asiakkaasta tarvittava vähimmäistietomäärä), RAI-käsikirjasta (sisältää muun muassa hoidon suunnitteluohjeiston) ja mittareista (asiakasrakenne, laatuindikaattorit jne.).

Osastojen ja laitosten henkilöstö hyödyntävät arviointitietoa esimerkiksi asiakkaan hoidon suunnittelussa, hoidon tarpeen arvioimisessa, voimavarojen jaossa ja laadun seurannassa. Kertyneistä tiedoista on luotu Stakesiin vertailutietokanta, jota osallistuvat yksiköt voivat käyttää internetin välityksellä kehittäessään omaa toimintaansa. Vertailukehittämisellä eli benchmarking-toiminnalla tarkoitetaan yleisesti toiminnan vertailua yksiköiden välillä ja parhaimmista käytännöistä oppimista. Hankkeessa kerätään vuosittain myös osastojen toimintatietoja, joita käytetään muun muassa yksikköjen tuottavuuden ja henkilöstömitoituksen laskennassa.

Tässä tutkimuksessa käytettävät osastojen toimintatiedot on kerätty sekä yhteistyössä laitosten talousvastaavien kanssa (talousjohtajat ja -sihteerit) että suoraan laitosten osastonhoitajilta ja osastojen työntekijöiltä. Lisäksi tietoja (esimer-

kiksi kliiniset laatuindikaattorit) on kerätty Stakesin ylläpitämästä RAI-hankkeen internetpohjaisesta vertailutietokannasta, joka muodostuu kaikista yksiköissä tehdystä asiakaskohtaisista RAI-arvioinneista.

8.2 Havaintoyksiköt

Osajulkaisussa 1 osastoja oli mukana 114 ja osajulkaisussa 2 osastoja oli 122. Osajulkaisussa 3 tutkimusjoukon muodosti 113 osastoa. Osajulkaisussa 4 käytettiin hierarkkista tutkimusmenetelmää, jossa havaintoyksikkönä olivat osastojen lisäksi myös asiakkaat. Aineisto koostui 1 787 asiakkaasta ja 68 osastosta. Osastot sijoituivat 16 eri laitokseen. Havaintoyksiköiden määrä vaihteli eri osatutkimuksissa, koska kaikkia eri osatutkimuksissa tarvittavia tietoja ei ollut saatavissa samoista yksiköistä. Lisäksi tutkimuksessa pyrittiin käyttämään aina uusinta käytettävissä olevaa aineistoa, mistä syystä aineistot ovat osin eri vuosilta. Tutkimuksessa käytetyt aineistot, menetelmät, analyysien viimevaiheessa käytetyt muuttujat on esitetty osajulkaisuittain taulukossa 2 (s. 56).

Useissa aiemmissa tutkimuksissa havaintoyksikkönä on ollut laitos (ks. luku 6.2). Tässä tutkimuksessa havaintoyksikköinä olivat kuitenkin yksittäiset osastot. Tämä turvasi riittävän havaintomäärän ja mahdollisti tulosten arvioimisen lähempänä mikrotason päätöksentekijöitä, sillä hoitoon liittyvät päätökset tehdään osastotasolla. Lisäksi osastotason analyysissa voitiin ottaa huomioon myös se, että sekä toiminnan laatu että tuotannollinen tehokkuus voivat vaihdella merkittävästi saman laitoksen eri osastoilla. Tässä tutkimuksessa otetaan siten aiempaa paremmin huomioon toiminnan eri tasot.

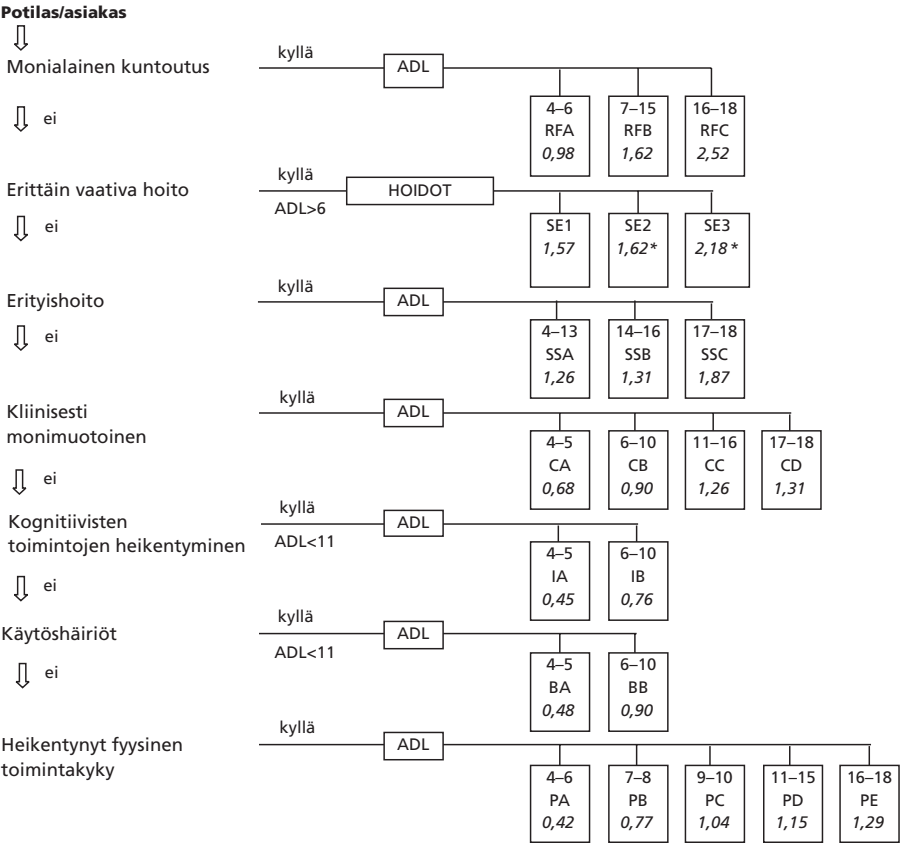
TAULUKKO 2. Osajulkaisuissa käytetyt aineistot, muuttujat ja menetelmät

	Aineisto	Muuttujat	Menetelmät
Osajulkaisu 1 <i>The Association between Quality of Care and Technical Efficiency in Long-Term Care</i>	114 vanhainkoti- ja terveyskeskus-osastoa Vuosi 2002	Panokset: kolmen ammattiryhmän henkilölukumäärät ja pääoma Tuotos: asiakasrakenteella vakioidut hoitopäivät (taloushallinnon ilmoittamat hoitopäivät) Laatuindikaattorit: 41 voimavaralaatua ja kliinistä prosessi- ja tuloslaatua kuvaavaa indikaattoria	DEA (panosorientoitunut vakioskaalatuoottomalli) ja Mann-Whitneyn U-testi
Osajulkaisu 2 <i>Measuring the Productive Efficiency and Clinical Quality of Institutional Long-Term Care for the Elderly</i>	122 vanhainkoti- ja terveyskeskus-osastoa Vuosi 2001	Tuotantofunktio: Panokset: kolmen ammattiryhmän henkilötövuodet ja pääoma (Box-Cox muunnos, $\lambda = 0,38$) Tuotos: asiakasrakenteella vakioidut hoitopäivät (Box-Cox muunnos $\theta = 0,22$, hoitopäiväkertymä laskettu asiakastasolta) Tehottomuuseffektit: Kolme kliinistä prosessi- ja tuloslaatuindikaattoria, osaston koko, laitostyyppi, kuoritusaste, osaston erikoistuminen ja asiakkaiden keski-ikä	Tuotantofunktio (yksivaiheisesti estimoitu stokastinen rintamalli)
Osajulkaisu 3 <i>The Cost Efficiency and Clinical Quality of Institutional Long-term Care for the Elderly</i>	113 vanhainkoti- ja terveyskeskus-osastoa Vuosi 2001	Kustannusfunktio: Kustannukset: toiminnan kokonaiskustannukset, € (logaritimuunnos) Tuotos: asiakasrakenteella vakioidut hoitopäivät (logaritimuunnos, hoitopäiväkertymä laskettu asiakastasolta) ja kaksi kliinistä tuloslaatuindikaattoria Hintamuuttuja: keskimääräisen henkilötövuoden hinta osastossa (logaritimuunnos) Toimintaympäristömuuttuja: laitostyyppi (vanhainkoti tai terveyskeskus) Tehottomuuseffektit: Kaksi kliinistä prosessilaatuindikaattoria	Kustannusfunktio (yksivaiheisesti estimoitu stokastinen rintamalli)
Osajulkaisu 4 <i>Patient- and ward-level determinants of nursing time in nursing facilities</i>	1 787 asiakasta, 68 palvelutalo-, vanhainkoti- ja terveyskeskus-osastoa ja 16 laitosta Vuosi 2002	Selitettävä: Asiakkaan palkkapainotettu hoitoaika, min/vrk Selittävät: RUG-III/22-luokitus, kipu, masennus, hoitajien antama kuntoutus, ikä, sisäänkirjauspäivämäärä, hoitotyön osuus, yksikön hoitopaikkamäärä, lääkitys, yhden hengen huoneiden osuus, moduulityöskentely, erikoistuminen, laitostyyppi ja huoneet, joissa on oma wc	Monitasomalli (asiakas- ja osastotasot) ja pienimmän neliösumman regressioanalyysi

8.3 Asiakasrakenne ja voimavaravaativuus

Osastojen asiakasrakenteen vakioimiseen käytettiin Yhdysvalloissa 1980-luvulla kehittyä RUG-luokitusjärjestelmää (Resource Utilization Groups), jonka avulla voidaan arvioida asiakkaiden vaatimia taloudellisia voimavaroja ja voimavaratarvetta (Fries, Cooney, & Leo, 1985; Fries et al., 1995). Luokitus on laajimmalle levinnyt asiakasrakennemittari vanhustenhuollossa ja se on käytössä maailmanlaajuisesti useassa maassa (Carpenter, Perry, Challis, & Hope, 2003; Carrillo et al., 1996; Fries et al., 1995; Ikegami et al., 1994). RUG-luokituksen viimeisimmän version, RUG-III:n, pätevyys on testattu Suomessa tutkimuksessa, jossa oli mukana noin 2000 vanhainkotien ja terveyskeskusten vuodeosastojen asukasta (Björkgren, Häkkinen, & Finne-Soveri, 1998). Suomessa sovelletaan RUG-III/22-luokitusta. Siinä asiakas luokitellaan aluksi johonkin seitsemästä kliinisesti yhdenmukaisesta pääryhmästä, jotka ovat voimavaratarpeen mukaan järjestettynä *monialainen kuntoutus, erittäin vaativa hoito, erityishoito, kliinisesti monimuotoinen hoito, kognitiivisten toimintojen heikentyminen, käytöshäiriöt ja heikentynyt fyysinen toimintakyky*. Seuraavassa vaiheessa asiakkaat luokitellaan alaryhmiin fyysisen toimintakyvyn eli RUG-III-ADL -indeksin mukaan, jonka arvo vaihtelee 4 ja 18 välillä. Suomessa käytössä olevassa RUG-III/22-luokituksessa näitä alaryhmiä on 22 (Björkgren et al., 1998). Jokaiselle alaryhmälle on laskettu kustannuspaino, joka kuvaa ryhmien välisiä voimavaratarpeen suhteellisia eroja. Voimavaratarpeella tarkoitetaan sitä, kuinka paljon eri ammattiryhmät antavat vuorokauden aikana asiakkaalle hoitoaikaa (min/vrk). Tämän avulla voidaan arvioida asiakkaasta koituva kustannus. RUG-III/22-luokituksessa keskimääräisen asiakkaan kustannuspaino on 1, vähiten hoitoa vaativan 0,42 ja eniten hoitoa vaativan 2,52. Eniten ja vähiten hoitoa vaativien ryhmien ero voimavaravaativuudessa on näin ollen kuusinkertainen. Pääomakustannusten oletetaan jakautuvan tasaisesti kaikille asiakkaille, joten niitä ei huomioida luokituksessa. RUG-III/22-luokituksen seitsemän pääluokkaa, 22 alaryhmää ja alaryhmien kustannuspainot on esitetty kuviossa 7 (s. 58).

RUG-III/22-luokitusjärjestelmä



HUOM: Kursivoidut numerot ovat RUG-III/22-ryhmien suomalaisia kustannuspainoja
* arvioitu kustannuspaino

Lähde: Björkgren ym. 1998.

KUVIO 7. RUG-III/22-luokitusjärjestelmä ja suomalaiset kustannuspainot

8.4 Tuotoksen mittaaminen

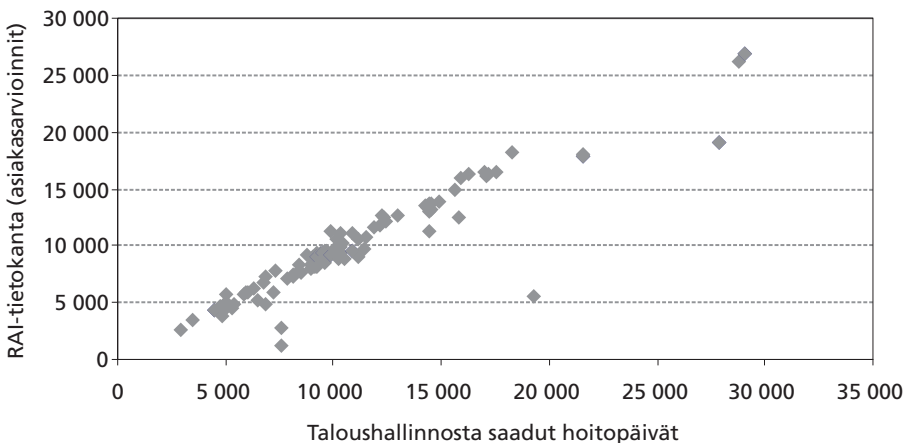
Tuotannollisen tehokkuuden mittaamisessa tuotosmittarina käytettiin osaston vuoden aikana tuottamia hoitopäiviä, jotka vakioitiin asiakasrakenteella. Osajulkaisussa 1 vakioidut hoitopäivät laskettiin talousyhdyshenkilöiden ilmoittamista hoitopäivistä, joita painotettiin osaston keskimääräisellä RUG-III/22-kustannuspainolla. Osaston keskimääräinen kustannuspaino laskettiin kahden RAI-arviointiaineiston keskiarvona. Nämä aineistot sisälsivät vuoden 2002 aikana osastolla tehdyt asiakaskohtaiset RAI-arvioinnit.

Osajulkaisuissa 2 ja 3 laskimme, kuinka monta päivää kukin vuoden 2001 aikana osastolla hoidossa ollut asiakas oli ollut kussakin RUG-III/22-ryhmässä. Näitä päiviä painotettiin kyseisen RUG-III/22-ryhmän kustannuspainolla. Lyhytaikaishoidossa olleita asiakkaita ei useinkaan ehditä arvioimaan ennen kotiutusta ja siirtämistä osastolta eteenpäin, joten lyhytaikaishoitopäivien kustannuspainoksi estimoititiin 1,42. Lyhytaikaisasiakkaat vaativat keskimäärin noin 40 prosenttia enemmän voimavaroja kuin pitkäaikaisasiakkaat. Kustannuspaino estimointiin käyttäen pienimmän neliösumman regressioanalyysia ja kustannuksia selitettiin pitkä- ja lyhytaikaishoitopäivillä. Osaston kokonaishoitopäiväkertymä vuoden aikana laskettiin summaamalla asiakkaiden painotetut lyhyt- ja pitkäaikaishoitopäivät eli

$$Y_k = \sum d_i w_i, \quad (6)$$

jossa d_i = kokonaishoitopäivät kussakin RUG-III/22-ryhmässä vuoden 2001 aikana ja w_i = RUG-III/22-ryhmän kustannuspaino.

Asiakasrakenteen vakioiminen RUG-III/22-luokituksen avulla voidaan tehdä tarkimmin yllä kuvatulla tavalla laskelmalla osaston RUG-III-kustannuspaino asiakaskohtaisista hoitopäiväkertymistä. Kuviossa 8 on havainnollistettu näitä kahden hoitopäivien laskentatapaa, jotka tuottavat hyvin yhtenevän tuloksen. Vuoden 2001 aineistossa asiakasasolta lasketun hoitopäiväkertymän korrelaatio taloushydyshenkilöiden ilmoittamiin hoitopäiviin oli 0,94. Poikkeavat osastot ovat akuutiosastoja, joiden osalta osajulkaisuissa 2 ja 3 käytettiin painotettuja taloushydyshenkilöiden ilmoittamia hoitopäiviä. Laskentamenetelmien tuottaman tuloksen yhtenevyyden vuoksi osajulkaisussa 1 käytettiin tuotosmittarina taloushydysh-



KUVIO 8. Taloushydyshenkilöiltä ja RAI-tietokannasta saatujen hoitopäivien välinen yhteys

henkilöiden ilmoittamia hoitopäiviä. Osastojen keskimääräinen RUG-III/22-kustannuspaino ei vaihtele vuosittain merkittävästi, vaikka laitoshoidossa olevien asiakkaiden keskimääräinen kustannuspaino on kohonnut viime vuosina (Noro, 2005).

8.5 Kustannustiedot ja panoshinnat

Osajulkaisussa 3 estimoidussa kustannusfunktioimallissa selitettävänä muuttujana käytettiin osaston vuoden 2001 kokonaiskustannuksia, jotka sisältävät myös pääomakustannukset ja poistot. Näin ollen kustannukset sisältävät kaikki toiminnasta yksikölle aiheutuvat kustannukset, jotka rahoitetaan julkisin varoin ja asiakasmaksuin. Koska tiedot kerättiin osastotasolta, laitoksien talousyhdyshenkilöitä pyydettiin kohdentamaan ja vyöryttämään kaikki laitoksen kustannukset osastoille. Riippuen kysytystä tiedosta vyörytys tehtiin joko hoitopäivien perusteella tai osaston pinta-alan mukaan. Pääasiassa vyörytys tehtiin hoitopäivien mukaan, mutta esimerkiksi toimitaloihin liittyvät välilliset kustannukset vyörytettiin osastojen pinta-alan mukaan.

Kustannusfunktioimallissa käytettiin yhtenä selittävänä muuttujana työvoiman hintaa. Panoshintana käytettiin keskimääräisen henkilötyövuoden hintaa osastossa. Tämä hinta laskettiin yksiköiden ilmoittamien palkkakustannusten perusteella. Palkkatietoa ei pystytty jakamaan eri ammattiryhmille. Tutkimuksessa käytettiin tästä syystä aggregoitua osastokohtaista työn hintamuuttujaa. Luotettavaa estimaattia pääoman hinnasta ei ollut saatavilla, joten erillistä pääomamuuttujaa ei käytetty. Tutkimuksessa jouduttiin oletamaan, että muut kuin työn hintaan liittyvät kustannukset ovat eri osastoilla suurin piirtein samansuuruisia (vrt. Linna & Häkkinen, 1999). Käyttämäämme hintamuuttujaa voidaan siten pitää estimaattina panoshinnoista ja siksi sitä kutsuttiin palkkahintamuuttujaksi.

8.6 Tuotannontekijät

Tuotannontekijöillä tarkoitetaan suoritteiden aikaansaamiseksi käytettäviä panoksia (ks. kuvio 3, s. 29). Osajulkaisussa 1 sovelletussa DEA-mallissa käytettiin panoksina osastolla jatkuvasti työskentelevien ammattiryhmien keskimääräisiä lukumääriä vuoden 2002 aikana. Henkilöstö jaettiin kolmeen ryhmään: osastonhoitajat ja sairaanhoitajat tai vastaavat, perus- ja lähihoitajat tai vastaavat ja avustava tai muu henkilöstö. Lisäksi panoksena käytettiin osaston hoitopaikkoja, joiden avulla mitattiin epäsuorasti pääomaa. Pääomaa ei yleensä pystytty julkisissa laitoksissa

sa mittaamaan suoraan, joten tavallisesti hoitopaikkojen määrää käytetään sijaismuuttujana (esim. Linna, 1999).

Osajulkaisussa 2 estimoidussa tuotantofunktiossa käytettiin puolestaan panosmuuttujina eri henkilöstöryhmien henkilötyövuosia vuodelta 2001. Lisäksi panoksena käytettiin osaston hoitopaikkoja, jolla mitattiin epäsuorasti pääomaa. Henkilöstö jaettiin kolmeen ryhmään: osastonhoitajat ja sairaanhoitajat tai vastaavat, perus- ja lähihoitajat tai vastaavat ja avustava tai muu henkilöstö. Henkilötyövuosilla tarkoitettiin säännöllistä normaalia vuosityöaikaa (252 työpäivää vuodessa). Vuosiloma, vapaapäivät ja sairaspoissaolot vähennettiin kuitenkin henkilötyövuosista. Henkilötyövuosia laskettaessa huomioitiin myös työntekijän koeaikaisuutta kuvaava työaika kerroin.

8.7 Aikamittaus

Marraskuussa 2002 Stakesin ja Suomen Kuntaliiton toteuttamassa aikamittaustutkimuksessa oli mukana 86 vanhainkoti-, terveyskeskus- ja palvelutalo-osastoa 19 laitoksesta Helsingistä, Espoosta ja Vantaalta. Näissä yksiköissä oli hoidossa aikamittausajankohtana 2 294 asiakasta. Aikamittausmenetelmänä käytettiin RUG-III-validointitutkimuksissa käytettyä menetelmää (Björkgren et al., 1999; Carpenter et al., 2003; Fries et al., 1985). Osaston hoitohenkilökunta (n = 1 031) kirjasi lomakkeelle työajan, joka kohdentui yleisiin töihin, jotka jaettiin viiteen pääryhmään, muun muassa hallinnollisiin ja ylläpitotöihin. Tämän lisäksi työntekijät merkitsivät asiakaskohtaisesti kullekin asiakkaalle suorasti ja epäsuorasti kohdennetun työajan. Asiakaskohtainen aika sisälsi kaikki sellaiset työtehtävät, jotka voitiin kohdentaa yksittäiselle asiakkaalle kuten esimerkiksi välitön hoitotyö, hoidon suunnittelu ja asiakkaan asiakirjojen täyttäminen. Työajan seuranta toteutettiin yhden tavanomaisen vuorokauden aikana (24h), mitä pidetään yleisesti riittävänä (Björkgren et al., 1999). Muu henkilöstö (n = 300), kuten esimerkiksi kuntoutushenkilöstö ja lääkärit kirjasivat vain asiakaskohtaisen työaikansa seitsemän vuorokauden ajalta. Omaisten (n = 1 167) antama hoitoaika kirjattiin myös seitsemän vuorokauden ajalta. Omaisten antamasta hoitoajasta huomioitiin vain sellaiset tehtävät, jotka korvasivat osastojen työntekijöiden työtä. Korvaavuuden arvioivat ja kirjasivat osaston työntekijät.

Osajulkaisussa 4 selitettävänä muuttujana käytettiin asiakkaan hoitoaikaa vuorokaudessa (min/vrk). Tämä hoitoaika laskettiin summana hoitohenkilöstön, muun henkilöstön ja omaisten merkitsemistä ajoista. Aikoja painotettiin vielä kunkin ammattiryhmän palkoilla, jolloin selitettävänä muuttujana voitiin käyttää asiakkaan palkkapainotettua hoitoaikaa (min/vrk). Palkkapainotettua työaikaa käytetään muun muassa RUG-III/22-järjestelmän alaluokkien kustannuspainojen

laskennassa, koska se kuvaa asiakaskohtaista kustannusta ja voimavaravaativuutta (Björkgren et al., 1999; Fries et al., 1985; Fries et al., 1995). Suhteelliset palkkapainot saatiin Tilastokeskuksen palkkatilastoista vuodelta 2001 (StatFin, 2004). Sairaanhoidajien palkkapaino oli 1, lähi- ja perushoitajien 0,89, lääkäreiden 1,99 jne. Omaisten ajalle annettiin painoksi 0,5, mutta myös 0,75 painoa testattiin.

8.8 Muut tiedot

Osajulkaisussa 2 ja 3 estimoiduissa malleissa arvioitiin usean osaston toimintaa ja rakennetta kuvaavan tekijän yhteyttä tuotannolliseen tehokkuuteen. Käytettävissä olevat muuttujat olivat laitostyyppi, osaston erikoistuminen psykiatria–dementia-asiakkaiden hoitoon, osastopalvelujen määrä, moduulityöskentely, asiakasvaihtuvuus, perus- ja lähihoitajien osuus henkilöstöstä, osaston erikoistuminen lyhytaikaisten asiakkaiden hoitoon, työkierron käyttäminen, osaston hoitopaikkojen määrä, kuormitusaste, osaston asiakkaiden keski-ikä ja osaston asiakkaiden sukupuolirakenne. Nämä tiedot saatiin osastonhoitajille suunnatusta lomakekyselystä ja RAI-tietokannasta. Useimpien muuttujien vaikutusta ei pystytty ennustamaan aiempien tutkimusten ja kirjallisuuden perusteella. Osa käytettävissä olevista muuttujista oli sellaisia, joita ei ole käytetty missään aiemmissa tuottavuustutkimuksissa. Björkgrenin (Björkgren, 2002) tulosten perusteella tässä tutkimuksessa voidaan kuitenkin odottaa, että pieni hoitopaikkamäärä, alhainen kuormitusaste ja perus- tai lähihoitajien pieni osuus koko henkilöstöstä saattavat olla yhteydessä tuotannolliseen tehostumukseen.

Osajulkaisussa 4 asiakkaalle kohdistetun hoitoajan varianssia selittävinä osastotason tekijöinä kokeiltiin yllä olevien tekijöiden lisäksi myös asiakkaille annettavien lääkkeiden keskimääräistä määrää osastolla (lääkkeet yhteensä ja erikseen antipsykootit, anksiolyytit, antidepressantit ja unilääkkeet), yhden hengen huoneiden osuutta, sellaisten huoneiden osuutta, joissa oli oma wc, sairaanhoitajien osuutta henkilöstöstä ja hoitotyön osuutta kokonaistyöajasta. Asiakkaiden ominaisuuksina kuvaavina tekijöinä käytettiin asiakkaan RUG-III/22-indeksien lisäksi erikseen fyysisistä toimintakykyä (seitsemänluokkainen ADL-mittari), kognition tasoa (seitsemänluokkainen kognitiotason mittari), kanssakäynnin ongelmia muiden kanssa, masennusoireita, päivittäistä kipua, osastolle tulopäivää, ikää ja hoitajien asiakkaalle antamaa kuntoutusta. Nämä tiedot saatiin RAI-tietokannasta (alun perin MDS 2.0-lomakkeesta).

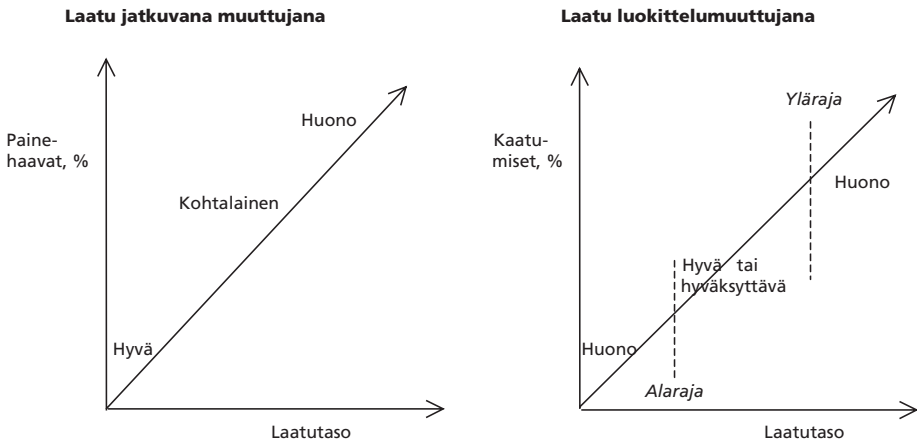
8.9 Laatumuuttajat

Tutkimuksessa pyrittiin käyttämään ensisijaisesti sellaisia laatua kuvaavia muuttujia, jotka ovat yksiköiden välillä vertailukelpoisia ja joita voidaan tutkitusti käyttää hoidon laadun indikaattorina ja jotka ovat saatavissa kaikista yksiköistä (ks. luku 4.4). Tästä syystä päädyttiin käyttämään pääasiassa RAI-tietokannasta saatavia hoidon kliinisiä laatuindikaattoreita. Hoidon kliinistä laatua mitattiin Yhdysvalloissa Wisconsinin yliopistossa kehitetyillä laatuindikaattoreilla. Niiden avulla voidaan arvioida haitallisten ja ei-toivottujen hoitoprosessien ja -tulosten yleisyyttä vanhusten laitoshoidossa (Karon & Zimmermann, 1996; Zimmermann et al., 1995). Tästä syystä kliinisissä laatuindikaattoreissa nimenomaan korkeampi laatuindikaattorin arvo (%) merkitsee heikompaa laatua. Nämä laatuindikaattorit ovat kattavasti saatavissa kaikista osastoista RAI-tietokannasta (Noro, 2005; Noro et al., 2001). Kaiken kaikkiaan kliinisiä laatualueita on toistakymmentä, joissa on kaiken kaikkiaan 41 laatuindikaattoria. Osa laatuindikaattoreista on riskivakioituja. Liitteessä 1 on esitetty laatualueet, yksittäiset laatuindikaattorit, indikaattorin tyyppi ja niiden laskenta MDS 2.0-lomakkeesta. Näistä indikaattoreista osajulkaisussa 2 käytettiin kolmea ja osajulkaisussa 3 neljää. Ne valittiin sisällöllisin perustein. Tämä tarkoittaa, että käytetyn indikaattorin tuli olla käyttökelpoinen sekä vanhainkodeissa että terveyskeskuksissa, indikaattoria on käytetty aiemmin laadun arvioinnissa eikä indikaattoria saanut korreloida voimakkaasti muiden muuttujien kanssa. Osajulkaisussa 2 ja 3 käytettyjen laatuindikaattorien määrä oli hyvin rajoitettu, koska havaintoja oli kohtuullisen vähän suhteessa estimoitavien parametrien määrään (ks. taulukko 2, s. 56). Muuttujien määrää kontrolloimalla haluttiin turvata se, etteivät mallin vapausasteet muodostua ongelmaksi. Osajulkaisussa 1 käytettiin epäparametrista lähestymistapaa, joka mahdollisti kaikkien kliinisten laatuindikaattorien käyttämisen. Tämän lisäksi käytettiin vielä kolmea voimavara- ja rakennelaatua kuvaavaa muuttujaa, jotka olivat sairaanhoitajien osuus koko henkilöstöstä, yhden hengen huoneiden osuus ja sellaisten huoneiden osuus, joissa oli oma wc. Tutkimusta tehtäessä arvioitiin, että rakennelaatumuuttujien suuri prosenttiosuus merkitsisi hyvää laatua. Osajulkaisussa 1 käytetyt rakennelaatua kuvaavat muuttajat saatiin osastonhoitajille lähetetystä lomakekyselystä. Laatumuuttujien tilastolliset tunnusluvut on esitetty osajulkaisun 1 taulukossa 1

Osajulkaisussa 2 ja 3 osastojen laatuindikaattorit laskettiin kolmen RAI-aineiston keskiarvona (laatuindikaattorin % keskimäärin) ajalta 16.9.2000–15.4.2002. Aineistot muodostuvat asiakaskohtaisista arvioinneista, jotka lähetetään osastoilta Stakesiin puolivuositain. Osajulkaisussa 1 osastojen laatuindikaattorit laskettiin puolestaan kahden RAI-arvioinnin keskiarvona vuodelta 2002. Laatuindikaattorien laskeminen useasta poikkileikkausaineistosta vähentää satunnaisuuden ja mahdollisten mittausvirheiden vaikutusta tuloksiin. Yksittäisen laatuindikaattorin arvo vaihtelee luonnollisesti jonkin verran ajan mittaan, jos osastossa on esimerkiksi

käynnistetty laatuun liittyviä kehittämishankkeita tai interventioita. Tällainen kehittämishanke voisi olla esimerkiksi pyrkimys vähentää asiakkaiden painehaavojen määrää kehittämällä hoitokäytäntöjä. Tällöin usean poikkileikkausaineiston käyttäminen tasoittaa luonnollisesti vaihtelua.

Yhdysvalloissa klinisille laatuindikaattoreille on laadittu myös raja-arvot (Rantz et al., 1997). Niiden perustana on asiantuntijapaneelin käsitys siitä, millä välillä hyvä hoidon laatu tulisi vaihdella. Kuviossa 9 on havainnollistettu laadun käsitteellistämistä jatkuvana muuttujana ja raja-arvoihin perustuvana luokittelumuuttujana. Laadun käsitteleminen jatkuvana muuttujana merkitsee, että mitä useammin haitallisia laatuilmiöitä, esimerkiksi painehaavoja esiintyy, sitä huonompi on hoidon laatu. Luokittelumuuttujana laatu on puolestaan hyvää, jos laatuilmiön yleisyys on jollakin raja-arvojen määrittämällä välillä. Esimerkiksi kaatumisten osalta raja-arvojen käyttäminen voi olla perusteltua siitä syystä, että kaatumisia esiintyy harvoin, jos asiakkaita estetään liikkumasta käyttämällä erilaisia liikkumista estäviä välineitä. Kaatumisia voi puolestaan esiintyä tavallista enemmän, jos asiakkaita kannustetaan liikkumaan itsenäisesti omatoimisuuden ja toimintakyvyn ylläpitämiseksi, mikä ei kuitenkaan kerro huonosta hoidosta. Toisaalta kaatumisia voi esiintyä runsaasti, jos osastolla ei ole riittävästi henkilökuntaa, mikä puolestaan kertoo huonosta laadusta. Osajulkaisussa 2 testattiin Yhdysvalloissa käytettyjä raja-arvoja, joiden lisäksi muodostimme omat hyvän laadun raja-arvot kunkin laatuindikaattorin laatu jakauman välille 10–90 prosenttia ja 5–95 prosenttia. Hoitokulttuuri- ja -käytäntöerojen vuoksi osa Yhdysvalloissa käytetyistä raja-arvoista ei sovellu kovin hyvin Suomeen (Finne-Soveri, 2005). Osajulkaisussa 1 asetimme puolestaan huonon laadun raja-arvoksi kaikilla klinisillä laatuindikaattoreilla ylimmän kvartiilin (75–100 %) tai ylimmän 15 persentiilin (85–100 %). Useimmissa aiemmissa tuottavuustutkimuksissa laatu muuttujia on käytetty vain jatkuvina muuttujina.



KUVIO 9. Hoidon laatu jatkuvana ja luokittelumuuttujana

9 Tutkimusmenetelmät

9.1 Parametriset menetelmät

9.1.1 Stokastiset rintamamallit ja monitasomallinnus

Osajulkaisussa 2 estimointiin tuotantofunktio soveltamalla Battesen ja Coellin mallispesifikaatiota poikkileikkausaineistoon, jossa tuotantofunktio ja tehottomuusefektit estimoidaan yksivaiheisesti suurimman uskottavuuden menetelmän avulla. Mallit estimointiin FRONTIER 4.1-ohjelmistolla (Battese & Coelli, 1993; Coelli, 1996). Tehottomuusefektiin eli tekijöihin, jotka voivat vaikuttaa osaston tuotannolliseen tehokkuuteen, sisällytettiin yksitoista osaston toimintaa ja olosuh-teita sekä hoidon laatua kuvaavaa muuttujaa. Mallin tehottomuuskomponentti oli lineaarinen funktio näistä tekijöistä siten, että

$$\mu_i = \delta_0 + \delta_1 + \dots + \delta_{11} + w_i, \quad (7)$$

jossa δ -parametrit ovat estimoitavia tehottomuusefektejä ja w_i on virhetermi, joka noudattaa katkaistua normaalijakaumaa ja sen keskiarvo on 0 ja varianssi σ^2 kuitenkin siten, että μ_i on positiivinen.

Osajulkaisussa 3 estimointiin kustannusfunktio, jossa osa osaston piirteitä ja hoidon tulosta kuvaavista muuttujista sisällytettiin kustannusfunktioon. Kaikki toiminnan tuotosta kuvaavat muuttujat eli hoitopäivät ja tulosta indikaattorit olivat siten osa mallin kustannusfunktio-osiota eli kustannuksiin vaikuttavia tekijöitä. Kuormitusaste ja kaksi hoidon prosessilaatua kuvaavaa laatuindikaattoria sisällytettiin vuorostaan erillisiksi tehottomuusefekteiksi, jolloin ne olivat kustannustehokkuuteen vaikuttavia, toimintaprosessia kuvaavia tekijöitä. Tällaista mallinnusta toiminnan laadun osalta ei ole käsitykseni mukaan aiemmin sovellettu sosiaali- ja terveydenhuollon tuottavuustutkimuksissa. Käytetyn mallin avulla voidaan arvioida prosessi- ja tuloslaadun yhteyttä sekä suoraan kustannuksiin että kustannustehokkuuteen.

Osajulkaisussa 4, jossa tarkasteltiin asiakkaille kohdennetun työajan varianssia selittäviä tekijöitä, käytettiin pienimmän neliösumman regressioanalyysin lisäksi monitasomallinnusta. Monitasomallinnus on laajasti käytetty menetelmä silloin, kun tutkimusaineisto on hierarkkisesti rakentunut (Goldstein, 1995; Rice & Jones, 1997; Rice & Leyland, 1996). Monitasomallinnuksen avulla voidaan selittää muuttujassa oleva vaihtelu jakaa eri tasoille, esimerkiksi asiakas- ja osastotasolle. PNS-menetelmää käytettäessä oletetaan Gauss-Markov ehdon mukaisesti,

että havainnot ovat riippumattomia toisistaan (ks. Maddala, 1992), mikä ei päde, jos aineisto on hierarkkisesti rakentunut. On esimerkiksi mahdollista, että samalla osastolla hoidetut asiakkaat ovat samankaltaisempia ja tulevat kohdelluiksi samankaltaisemmin kuin eri osastoilla olevat asiakkaat. Parametriestimaattien keskivirheet ovatkin epätarkkoja, jos havainnot eivät ole toisistaan riippumattomia. Lineaarinen monitasomalli (tässä kaksi tasoa) voidaan esittää muodossa

$$y_{ij} = \alpha + \beta_p x_{ij} + (u_j + e_{ij}), \tag{8}$$

joka on normaali regressioyhtälö, mutta jossa virhetermi on jaettu siten, että u_j on korkeamman tason residuaali ja e_{ij} on alemman tason residuaali. Alaindeksi j viittaa esimerkiksi osastoon ja alaindeksi i asiakkaaseen. Virhetermit ovat satunnaismuuttujia, jotka ovat jakautuneet normaalisti ja niiden keskiarvo on 0 ja varianssi σ^2 . Osajulkaisussa 4 käytettiin kaksitasoista mallia, jossa selittävät muuttujat oli mitattu yksilö- ja osastotasolla. Mallien estimointiin käytettiin MLwiN 1.1-ohjelmistoa (Rasbash et al., 2001). Monitasomallin etu kiinteiden efektien malliin eli indikaattorimuuttujamalliin, jossa toisen tason havainnot ovat mukana dummymuuttujina, on esimerkiksi se, että monitasomalli säästää vapausasteita.

9.1.2 Malliin liittyvien oletusten ja mallispesifikaation testaaminen

Stokastiset tuotanto- ja kustannusfunktiot ovat herkkiä mallispesifikaatiolle ja malliin liittyville oletuksille, kuten esimerkiksi malliin valitulle muuttujakombinaatiolle, mallin funktionaaliselle muodolle ja tehottomuusefektien valinnalle (Battese & Broca, 1997; Maddala, 1992). Tästä syystä osajulkaisussa 2 ja 3 käytettyjä malleja arvioitiin usealla kriteerillä ja tilastollisella testillä.

Tuotanto- ja kustannusfunktioimallien funktionaalista muotoa tarkasteltiin Box-Cox-analyysin avulla. Box-Cox-analyysia käytetään yleisesti muun muassa mallin virhetermin jakauman normalisointiin ja riippuvuuksien lineaarisointiin (ks. Greene, 1993). Muunnos on muotoa $y^{(\lambda)} = (y^\lambda - 1) / \lambda$, ja se voidaan tehdä sekä selitettävälle että selittävälle muuttujille. Vaihtoehtoisia malleja eri muunnosparametrein, esimerkiksi lineaarista ja logaritmista mallia, verrattiin log-likelihood-testillä, joka on puolestaan muotoa $\lambda = -2[\ln(LH_0)] - [\ln(LH_1)]$. Osajulkaisussa 2 päädyttiin selitettävän muuttujan ja tuotantofunktion panosmuuttujien Box-Cox-muunnokseen. Tehottomuusefekteihin ei tehty muunnosta. Osajulkaisussa 3 estimointiin aluksi translog-malli, joka sisälsi kustannusfunktion tuotos- ja panoshintamuuttujat sekä alkuperäisessä muodossa, toisen asteen termeinä että ristitermeinä. Toisen asteen termit ja ristitermit eivät olleet tilastollisesti merkitseviä, joten mal-

li redusoitui logaritmiseksi kustannusfunktiomalliksi. Tulos- ja prosessilaatumuuttujiin ei kuitenkaan tehty muunnoksia.

Osajulkaisussa 2 ja 3 stokastisten tuotanto- ja kustannusfunktiomallien tehottomuuskomponentin spesifikaatiota arvioitiin log-likelihood-testillä, jonka avulla voidaan esimerkiksi arvioida (i) tarvitaanko stokastista rintamamallia eli onko $\text{var}(u) \neq 0$, vai voidaanko malli estimoida PNS-menetelmän avulla, (ii) noudattaako tehottomuusvirhetermi puolinormaalia vai katkaistua normaalijakaumaa⁶ ja (iii) tarvitaanko mallissa jotakin valittua tehottomuusefektimuuttujien kombinaatiota eli onko $\delta_1, \delta_2, \dots, \delta_n \neq 0$. Mikäli yksiköiden välinen vaihtelu ei selittyisi tehokkuuseroilla, voitaisiin kustannusfunktio estimoida pienimmän neliösumman regressioanalyysin avulla.

Odotusten mukaisesti jotkut osajulkaisuissa 2 ja 3 käytetyistä osastotason muuttujista korreloivat kohtuullisesti keskenään. Muuttujien väliset korrelaatiot saattavat aiheuttaa mallissa multikollineaarisuutta, jolloin parametriestimaatit saattavat olla epäluotettavia. FRONTIER 4.1-ohjelmistossa multikollineaarisuuden arvioimiseen ei ole testejä, joten sitä arvioitiin kokeilemalla eri muuttujakombinaatioiden vaikutusta mallien parametriestimaatteihin. Multikollineaarisuutta aiheuttavat muuttujat jätettiin malleista pois. Malleissa testattiin systemaattisesti myös muuttujien toisen asteen termit ja yhdysvaikutukset.

Osajulkaisussa 4 sovellettiin monitasomallinnusta, jossa ensimmäisenä tasona olivat asiakkaat ja toisena tasona osastot. Testasimme myös laitostason merkittävyyden, mutta malli redusoitui asiakas- ja osastotaseiseksi, koska kolmannella tasolla ei ollut tilastollisesti merkitsevää vaihtelua. Multikollineaarisuutta ja mallin oletusten toteutumista tarkasteltiin korrelaatioiden ja mallidiagnostiikan avulla. Muuttujien välillä ei ollut merkittäviä korrelaatioita. Mallidiagnostiikan perusteella lineaarinen malli todettiin käyttökelpoiseksi. Lisäksi tarkasteltiin poikkeavia ja vaikutusvaltaisia havaintoja asiakas- ja osastotasolla muun muassa yksittäisten havaintojen vetovoima-arvojen (lävistäjäalkion), Cookin etäisyyden (Cookin D:n) ja (standardoitujen) residuaalien avulla. Poikkeavilta näyttävien osastojen tai asiakkaiden poistaminen mallista ei vaikuttanut mallin estimaatteihin, joten osajulkaisussa voitiin käyttää kaikkia käytettävissä olevia asiakas- ja osastotason havain-toja.

⁶ FRONTIER 4.1-ohjelmistossa, jonka avulla mallit estimoitiin, ei voi käyttää tehottomuustermiin eksponentiaalista jakaumaa tai gammajakaumaa.

9.2 Epäparametriset menetelmät

9.2.1 Data envelopment -analyysi

Osajulkaisussa 1 sovellettiin DEA-menetelmää teknisen tehokkuuden laskemiseen. DEA-malli voidaan laskea niin, että siinä käytetään vakioskaalatuottojen (CRS) oletusta tai niin, että siinä käytetään muuttuvien skaalatuottojen (VRS) oletusta. Lisäksi laskenta voidaan tehdä joko panos- tai tuotosorientoituneesti. Vakioskaalatuottomallissa oletetaan, että yksiköiden tuotos ja käytetyt voimavarat kasvavat samassa suhteessa. Orientaatio ei vaikuta tuloksiin, jos laskennassa oletetaan vakioskaalatuotot. Jos oletuksena ovat muuttuvat skaalatuotot, niin panos- ja tuotosorientaatiot tuottavat lähtökohtaisesti erilaisen tuloksen. Päädyimme käyttämään panosorientoitunutta DEA-mallia, koska osastojen oletettiin pystyvän vaikuttamaan paremmin tuotannontekijöiden (panosten) kuin tuotettujen hoitopäivien määrään. Panoksina käytettiin kolmen eri ammattiryhmän keskimääräisiä lukumääriä osastolla vuoden aikana ja osaston hoitopaikkamäärää, jonka avulla mitattiin epäsuorasti pääomaa. Laskimme DEA-mallin sekä vakioskaalatuotto-oletuksella että muuttuvien skaalatuottojen oletuksella. Nämä kuitenkin tuottivat suhteellisen yhdenmukaiset tehokkuusluvut yksiköille. Jatkoanalyysissä käytettiin vain panosorientoituneen vakioskaalatuottomallin tuottamia teknisen tehokkuuden lukuja.

Laskettaessa teknistä tehokkuutta, merkitään yksikön tuotosvektoria m -ulotteisella vektorilla $y = (y_1, \dots, y_m)$ ja panoksia k -ulotteisella vektorilla $x = (x_1, \dots, x_k)$. Tekninen tehokkuus kullekin yksikölle (μ) saadaan maksimoimalla painotettujen tuotosten ja panosten suhde. Painot määräytyvät siten, että ne ovat tarkasteltavalle yksikölle suotuisimmat. Painot ja yksikön tehokkuusluku voidaan ratkaista lineaarisena optimointiongelmana seuraavasti:

$$\begin{aligned} \text{Min}_{z, \mu} \quad & \mu \\ \text{s.t.} \quad & z \cdot Y \geq y_i, \\ & z \cdot X \leq \mu \cdot x_i, \\ & z_i \geq 0 \end{aligned} \tag{9}$$

missä Y on $m \times k$ tuotosmatriisi, X on $n \times k$ panosmatriisi jokaiselle havaintoyksikölle, alaindeksi i on tarkasteltava yksikkö ja z on painokerroinvektori. Vakioskaalatuotto-oletuksen mukaan painojen summa rajoitetaan positiiviseksi, kun taas muuttuvien skaalatuottojen teknologiassa painojen summa rajoitettaisiin yhdeksi ($z_i = 1$). Data envelopment -analyysissä on myös mahdollista rajoittaa panosten ja tuotosten painoja vapaasti esimerkiksi tuotosten yksikköhintojen perusteella

(esim. Aaltonen et al., 2004). Tällä voidaan yrittää estää ei-toivottuja painoja, kuten esimerkiksi nollapaino jollekin tuotos-panos-suhteelle. Tässä tutkimuksessa painoille ei asetettu erityisiä rajoitteita.

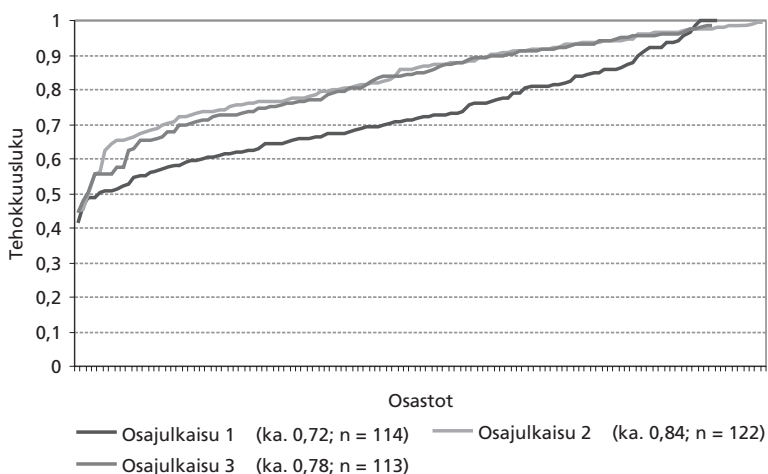
9.2.2 Hyvän ja heikon hoidon laadun osastojen vertailu tuotannollisen tehokkuuden suhteen

Osajulkaisussa 1 käytettiin Mann-Whitneyn U-testiä, joka on järjestyslukuihin perustuva epäparametrinen, kahden jakauman sijainnin vertailumenetelmä. Menetelmä ei vaadi oletuksia jakaumasta kuten parametriset menetelmät (esim. t-testi). Ensimmäisessä vaiheessa havaintojoukon muodostaneet 114 osastoa asetettiin paremmuusjärjestykseen erikseen kunkin laatuindikaattorin mukaan. Tämän jälkeen muodostettiin laadun raja-arvokriteereillä hyvän ja huonon laadun osastojen ryhmät ja verrattiin, eroavatko ryhmien teknisen tehokkuuden luvut (järjestys) tilastollisesti merkitsevästi toisistaan. Käyttämämme menetelmä on robusti verrattuna parametrisiin menetelmiin ja se johtaa asymptoottisesti normaaleihin (khi) jakaumiin Mann-Whitneyn U-testisuurelle. Epäparametrista menetelmää käyttämällä pyrimme myös välttämään kaksivaiheiseen lähestymistapaan liittyviä ongelmia, jotka koskevat erityisesti parametriestimaattien harhaisuutta (ks. luku 8.1). Epäparametrinen lähestymistapa mahdollistaa myös kaikkien käytettävissä olevien laatuindikaattorien käytön, koska tällöin muuttujien väliset korrelaatiot (multikolinearisuus) tai mallin vapausasteet eivät ole ongelmallisia.

10 Tulokset

10.1 Pitkäaikaishoidon tuotannollinen tehokkuus (osajulkaisut 1, 2 ja 3)

Tutkimuksessa havaittiin, että pitkäaikaishoidon tuotannollisessa tehokkuudessa on merkittäviä eroja. Osajulkaisussa 1 osastojen tekninen tehokkuus laskettiin käyttämällä DEA-menetelmää ja aineistoa vuodelta 2002. Keskimääräiseksi tehottomuudeksi saatiin 28 prosenttia, joka on hieman suurempi kuin aiemmissa tutkimuksissa. Osajulkaisussa 2 arvioitiin osastojen teknistä tehokkuutta estimoimalla tuotantofunktio käyttäen stokastista rintamamallia. Keskimääräinen tehottomuus estimoidun mallin mukaan oli vuonna 2001 noin 16 prosenttia. Osajulkaisussa 3 estimointiin kustannusfunktio käyttäen stokastista rintamamallia. Aineistona käytettiin myös vuoden 2001 aineistoa. Keskimääräiseksi kustannustehottomuudeksi saatiin 22 prosenttia. Eri osajulkaisujen perusteella voidaan arvioida, että tuotannollinen tehottomuus laitoshoidossa oli keskimäärin noin 20 prosenttia. Tämä tarkoittaa, että vanhainkotien ja terveyskeskussairaaloiden vuodeosastot käyttivät keskimäärin viidenneksen enemmän henkilöstö- ja pääomavaroja hoitopäivien tuottamiseen kuin kaikkein tehokkaimmat yksiköt, jotka olivat vertailuyksiköinä. Tutkimuksen osajulkaisuissa 1–3 saadut tehokkuuslukujen jakaumat, keskiarvot ja havaintomäärät on esitetty kuviossa 10. Tehokkuusluvun arvo 1 tarkoittaa tehokasta ja alle yhden oleva arvo tehottomuutta.



KUVIO 10. Tehokkuuslukujen jakaumat osajulkaisuissa 1–3

Tehokkuusluvut ovat aina aineisto- ja menetelmäsidoonaisia, joten ne eivät ole vertailukelpoisia keskenään. Eri menetelmiin ja aineistoihin perustuvia tehokkuuslukuja ei voi siten koskaan suoraan verrata toisiinsa. Tutkimus antaa kuitenkin suhteellisen yhdenmukaisen kuvan laitoshoidon keskimääräisestä tuotannollisesta tehottomuudesta ja sen jakaumasta. Pitkäaikaishoidon tuotannollinen tehottomuus ja sen vaihtelu on suuruudeltaan nykyään keskimäärin samaa luokkaa kuin 1990-luvun puolivälissä (Björkgren, 2002).

10.2 Hoidon laadun yhteys tekniseen tehokkuuteen (osajulkaisut 1 ja 2)

Osajulkaisussa 1 tarkasteltiin 38 kliinisen laatuindikaattorin ja kolmen rakenne-laatua kuvaavan indikaattorin yhteyttä tekniseen tehokkuuteen. Tutkimuksessa kiinnitettiin huomiota erityisesti osastoihin, joissa laatuindikaattoreina käytetyt muuttujat osoittivat selvimmin todennäköisiä laatuongelmia. Kolmentoista laatuindikaattorin osalta hoidon heikko laatu oli yhteydessä korkeaan tekniseen tehokkuuteen. Sitä vastoin kuusi heikkoa laatua kuvaavaa indikaattoria olivat yhteydessä tekniseen tehottomuuteen. 22 indikaattorin osalta yhteydet eivät olleet tilastollisesti merkitseviä. Tekniseen tehokkuuteen olivat pääasiassa yhteydessä sellaiset laatuominaisuudet, jotka kuvaavat passivoivaa ja ei-kuntouttavaa hoitotyötä tai tällaisen hoitotyön ja -käytännön tuloksia. Esimerkiksi osastot, joilla esiintyi runsaimmin vuoteeseen hoidettavia asiakkaita, liikkumista estävien välineiden käyttöä, painehaavoja, virtsan/ulosteen pidätyskyvyttömyyttä tai aktiviteettien vähäisyyttä olivat teknisesti tehokkaita. Rakenne- tai voimavaralaatua kuvaavista muuttujista sairaanhoitajien suuri osuus henkilöstöstä ja yhden hengen huoneiden suuri osuus kaikista potilashuoneista olivat yhteydessä tekniseen tehottomuuteen. Tulokset olivat myös suhteellisen samansuuntaisia tarkasteltuna erikseen vanhainkotiaineiston ja terveyskeskusaineiston perusteella. Tulokset eivät olleet yhteydessä laitostyyppiin. Taulukossa 3 (s. 72) on esitetty ne laatuindikaattorit, joiden perusteella näihin johtopäätöksiin päädyttiin. Taulukossa tulokset on esitetty sekä tehokkuuden järjestyslukuihin että tehokkuusluvun keskiarvoon perustuen. Osastot, joissa laatuindikaattorit osoittivat heikointa hoidon laatua, sijoituivat 114 osaston joukossa keskimäärin 25 sijaa korkeammalle tehokkuusluvun järjestyksen suhteen kuin hyvän laadun osastot. Tämä tarkoittaa käytännössä sitä, että huonon laadun osastoissa tehokkuusluku oli keskimäärin noin 0,11 yksikköä korkeampi kuin hyvän laadun verrokkiosastoissa.⁷

⁷ Myöhemmin tehdyn analyysin tulokset olivat samanlaiset kuin osajulkaisussa 1 esitetyt, vaikka asetelma oli päinvastainen eli kun osastot jaettiin tehokkuusluvun mukaan kahteen ryhmään, joiden välillä laatua verrattiin.

TAULUKKO 3. Teknisen tehokkuuden ero hyvän ja huonon hoidon laadun osastoissa¹

Laatuindikaattori	Mann-Whitneyn U-testi Ero osastojen teknisen tehokkuuden järjestys- luvuissa huonon (n = 17) ja hyvän (n = 97) laadun osastoryhmien välillä	t-testi Hyvän (n = 97) ja huo- non (n = 17) laadun osastoryhmien tehok- kuuslukujen keskiarvo- jen välinen ero
Kliiniset laatuindikaattorit		
Masennus ilman hoitoa, %	17	+0,08
Virtsan tai ulosteen pidätyskyvyttömyys, % alhainen riski	41	+0,18
Virtsan tai ulosteen pidätyskyvyttömyys, % korkea riski	29	+0,13
Katetri, % alhainen riski	23	+0,09
Katetri, % korkea riski	30	+0,13
Ulosteen pakkautuminen, %	17	+0,07
Vuoteeseen hoidettavat asiakkaat, %	41	+0,18
Harjoittelun/taitojen parantamisen/liike- laajuuden ylläpitämisen puute liikunta- kyvyttömillä asukkailla, %	17	+0,07
Liikkumista estävien välineiden päivittäi- nen käyttö, %	29	+0,11
Vähäinen/olematon aktiviteetteihin osallistuminen, %	25	+0,11
1–4 asteen painehaavat, % alhainen riski	24	+0,11
Voimavara- ja rakennelaatuindikaattorit		
Sairaanhoitajien osuus henkilöstöstä, %	21	-0,08
Yhden hengen huoneiden osuus kaikista asiakkaiden huoneista, %	15	-0,06

¹ Osajulkaisussa 1 ryhmien vertailu perustui järjestyslukuihin, mikä ei kuitenkaan havainnollisesti kuvaa eroa itse tehokkuusluvussa (teknisesti tehokas 1, heton < 1, keskiarvo 0,72). Esimerkiksi yhden sijan ero järjestyksessä voi perustua vain merkityksettömään 0,001 yksikön eroon tehokkuusluvussa. Järjestysluvuissa osastot joudutaan laittamaan paremmuusjärjestykseen, jolloin erojen absoluuttista suuruutta ei voida suoraan ottaa huomioon. T-testillä voidaan järjestyslukuja havainnollisemmin arvioida eroa tehokkuusluvuissa kahden ryhmän välillä, vaikka tässä esitettävissä t-testeissä käytetyt muuttujat eivät välttämättä täyttäneetkään t-testille vaadittua oletusta vasten normaalijakamasta. Osastojen kokonaismäärä oli 114.

Osajulkaisussa 2 käytettiin kolmea hoidon prosesseja ja tuloksia kuvaavaa kliinistä laatuindikaattoria ja tarkasteltiin niiden yhteyttä tekniseen tehokkuuteen. Masennus ja rauhoittavien lääkkeiden viikoittainen käyttö (%) ja masennus ilman hoitoa (%) olivat yhteydessä tehottomuuteen, mutta niiden osalta tilastollinen näyt-

tö (tilastollinen merkitsevyystaso) oli kuitenkin heikko. Sen sijaan painehaavojen yleisyys oli voimakkaasti ja tilastollisesti merkitsevästi yhteydessä tekniseen tehokkuuteen. Osastot, joilla oli paljon painehaavoista kärsiviä asiakkaita toimivat teknisesti tehokkaammin. Tulos oli tältä osin yhdenmukainen osajulkaisun 1 tulosten kanssa.

10.3 Hoidon laadun yhteys kustannustehokkuuteen ja kustannuksiin (osajulkaisu 3)

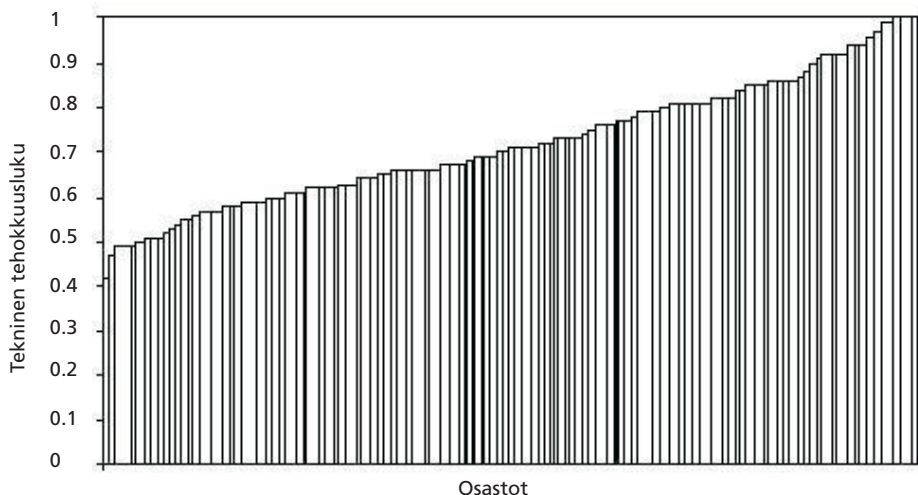
Osajulkaisussa 3 haitallisia hoitoprosesseja kuvaavat laatuindikaattorit, joita olivat masennuslääkkeiden sekä rauhoittavien lääkkeiden viikoittainen käyttö (%) ja liikkumista estävien välineiden käyttö (%) määriteltiin tehokkuuteen vaikuttaviksi prosessitekijöiksi. Sen sijaan hoitamattoman masennuksen (%) ja painehaavojen yleisyys (riskivakioitu, %), jotka kuvaavat haitallisia hoitotuloksia määriteltiin hoitopäivien tapaan toiminnan tuotoksiksi ja siten suoraan kustannusfunktion sisällytettäväksi tekijöiksi. Painehaavojen yleisyys lisäsi odotusten mukaisesti kustannuksia, mutta hoitamattoman masennuksen osalta tulos ei ollut tilastollisesti merkitsevä. Painehaavoille ja myös hoitamattoman masennuksen yleisyydelle estimoitiin joustoksi 0,02. Tämä tarkoittaa, että painehaavojen tai hoitamattoman masennuksen lisääntyminen 10 prosentilla lisää osaston kokonaiskustannuksia 2 prosenttia. Prosessilaatuindikaattoreista masennuslääkkeiden sekä rauhoittavien lääkkeiden viikoittainen käyttö lisäsi kustannustehottomuutta tilastollisesti merkitsevästi. Kustannustehokkuus aleni 0,13 yksikköä (tehokkuusluvun keskiarvo oli 1,22), kun lääkkeiden viikoittainen käyttö lisääntyi 35 prosentista 83 prosenttiin. Toisin sanoen kun siirryttiin kyseisen laatuindikaattorin jakauman alimmasta kymmenyksestä ylimpään kymmenykseen.

Kustannusfunktion estimointi tuotti suhteellisen samansuuntaisen tuloksen kuin osajulkaisussa 2 estimoitu tuotantofunktio. Yhteenvetona osajulkaisuista 2 ja 3 voidaan todeta, että hoidon huono laatu on yhteydessä sekä korkeisiin kustannuksiin että korkeaan tekniseen tehokkuuteen. Laadun vaikutus ei kuitenkaan ole erityisen voimakas ainakaan pienillä laatutason muutoksilla ja muiden tekijöiden pysyessä muuttumattomina. Toisin sanoen pieni muutos jollakin hoidon laadun osa-alueella suuntaan tai toiseen ei vaikuta merkittävästi tuotannolliseen tehokkuuteen tai kustannuksiin eikä päinvastoin. Tämä on yhdenmukainen tulos aiempien tutkimusten kanssa (esim. Knox et al., 2003; Rosko et al., 1995).

10.4 Muita tuotannolliseen tehokkuuteen yhteydessä olevia tekijöitä (osajulkaisut 2 ja 3)

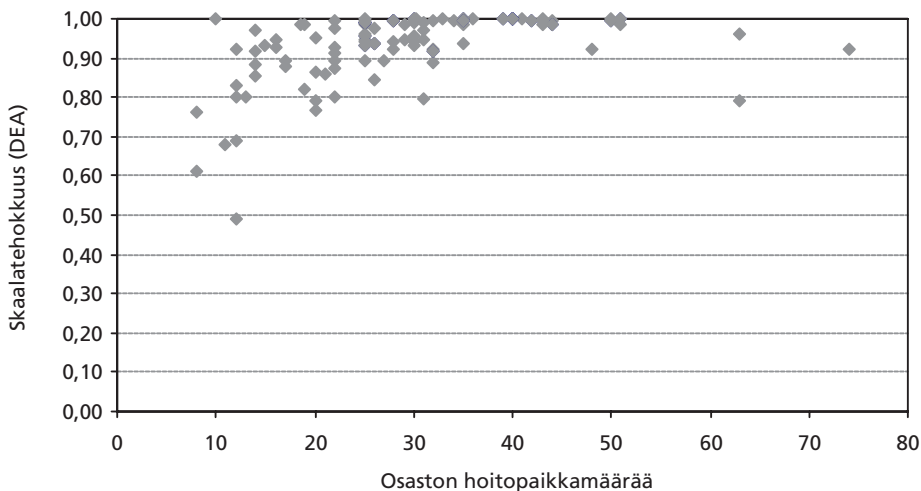
Laadun lisäksi tutkimuksessa testattiin useiden osaston toimintaa ja rakennetta kuvaavien muuttujien yhteyttä tuotannolliseen tehokkuuteen. Osastot, jotka olivat keskittyneet sekalaisten asiakkaiden hoitoon, olivat teknisesti tehokkaampia kuin psykiatristen ja dementia-asiakkaiden hoitoon erikoituneet osastot. Näin ollen pelkän teknisen tehokkuuden näkökulmasta asiakkaita tulisikin hoitaa sekaosastoilla. Lisäksi havaittiin, että terveyskeskusosastot tai osastot, joilla oli korkea kuormitusaste, olivat myös teknisesti tehokkaita.

Yksikön hoitopaikkojen määrä ei ollut yksiselitteisesti yhteydessä tekniseen tehokkuuteen. Kuviossa 11 on havainnollistettu Salter-diagrammin avulla yksikön koon ja teknisen tehokkuuden yhteyttä 114 osastossa. Kuviossa pylvään leveys kuvaa yksikön hoitopaikkojen määrää ja pylvään korkeus teknistä tehokkuutta. Hoitopaikkojen määrä vaihtelee 8:sta 74:ään, ja tehokkaat yksiköt saavat arvon 1. Teknisen tehokkuuden ja osaston koon välillä ei ole havaittavissa selvää yhteyttä. Sen sijaan mittakaavaedun eli niin sanotun skaalatehokkuuden ja yksikön hoitopaikkojen määrän välillä havaitaan samansuuntainen tulos kuin Björkgrenin tutkimuksessa (Björkgren et al., 2001). Mittakaavaetu on todennäköisesti heikoin hoitopaikkamäärältään keskimääräistä pienemmissä osastoissa, vaikka osa niistäkin toimii mittakaavaedun suhteen optimaalisesti (kuvio 12).⁸



KUVIO 11. Teknisen tehokkuuden ja osaston hoitopaikkojen määrän välinen yhteys

⁸ Kuvioissa 11 ja 12 suurimmat eli noin 50 paikkaa sisältävät osastot on muodostettu yhdistämällä pienempiä osastoja, koska kaikkia tehokkuuden laskemisessa tarvittavia tietoja ei ollut saatavissa yksittäisen osaston tasolta. Keskimääräinen osastokoko on laitoshoidossa 28–30 hoitopaikkaa.



KUVIO 12. Mittakaavaedun (skaalatehokkuus) ja osaston hoitopaikkojen määrän välinen yhteys

Tutkimuksessa osoitettiin lisäksi, että terveyskeskuksissa yksikkökustannukset olivat korkeammat kuin vanhainkodeissa, mikä ilmeisesti johtuu osittain kalliimmasta ja koulutetummasta henkilöstöstä tai korkeammista pääomakustannuksista. Kustannustehokkuutta selitettiin alustavissa malleissa samoilla osaston toimintaa ja luonnetta kuvaavilla tekijöillä kuin teknistä tehokkuutta osajulkaisussa 2. Näitä tekijöitä olivat muun muassa kuormitusaste, tiimi- ja moduulityöskentely ja osaston erikoistuminen. Kuormitusastetta lukuun ottamatta osaston toimintaympäristöä kuvaavat muuttujat eivät olleet yhteydessä kustannustehokkuuteen. Korkea kuormitusaste lisäsi odotusten mukaisesti osaston tehokkuutta. Toisin sanoen hoitopaikkojen tyhjäkäyttö aiheuttaa osastoissa kustannustehottomuutta. Tämä muuttuja jätettiin kuitenkin lopullisesta mallispesifikaatiosta pois (ks. taulukko 3 osajulkaisussa 3), koska tutkimuksen tarkoitus olisi ensisijaisesti tarkastella laadun yhteyttä kustannustehokkuuteen. Kuormitusasteen poistaminen mallista ei vaikuttanut merkittävästi muihin parametriestimaatteihin.

10.5 Työajan käyttö osastoissa (osajulkaisu 4)

Keskimääräinen asiakaskohtainen eri ammattiryhmien palkoilla painotettu hoitoaika vuorokaudessa oli 99 minuuttia vuonna 2002, joka on vajaa 20 minuuttia enemmän kuin vuonna 1995, jolloin aikamittaustutkimus toteutettiin edellisen kerran (Björkgren et al., 1998). Painottamattoman hoitoajan keskiarvo oli 111 minuuttia vuonna 2002. Palkkapainotetun ja painottamattoman ajan korrelaatio asiakastasolla oli 0,98. Yksittäisten asiakkaiden kohdalla hoitoaika vuorokautta kohden

TAULUKKO 4. Yksittäiselle asiakkaalle kohdentunut työaika ja työajan muutos vuodesta 1995 vuoteen 2002

	Vuosi 1995	Vuosi 2002	Muutos vuosien välillä, %
Asiakaskohtaisen työajan osuus kokonaistyöajasta, %	40	49	+ 18
Asiakkaalle annettu asiakaskohtainen hoitoaika, painottamaton min/vrk	83	111	+ 23
Asiakkaalle annettu asiakaskohtainen hoitoaika, palkkapainotettu min/vrk	83	99	+ 16

vaihteli 2 ja 424 minuutin välillä. Hoitoaikojen osastokeskiarvot vaihtelivat 66 ja 137 minuutin välillä. Osastolla työskentelevän henkilöstön työajasta 49 prosenttia kohdistui suorasti tai epäsuorasti yksittäisiin asiakkaisiin. Vuonna 1995 näin määritellyn hoitotyön osuus oli vain 40 prosenttia (Björkgren et al., 1999). Taulukossa 4 on esitetty työajan käytössä tapahtuneet keskeiset muutokset vuodesta 1995 vuoteen 2002. Muutokseen ovat ilmeisesti syynä asiakasrakenteen muuttuminen ja uudet hoitokäytännöt kuten esimerkiksi kuntouttavan työotteen lisääntyminen.

10.6 Työajan kohdentumista selittävät tekijät (osajulkaisu 4)

Tulosten mukaan asiakkaiden hoitoaika määräytyi ensisijaisesti asiakkaan ominaisuuksien perusteella. Näitä ominaisuuksia mitattiin RUG-III/22-luokituksella ja eräillä muilla MDS 2.0-lomakkeen tiedoilla. Asiakkaan RUG-III/22-indeksi perustuu asiakkaan useaan ominaisuuteen eli käytännössä kymmeneen MDS 2.0-lomakkeen tietoihin kuten esimerkiksi fyysiseen toimintakykyyn, kognitioon, terveydentilaan ja mielialaan. Asiakaskohtaisista tekijöistä hoitoaikaa lisäsivät voimakkaimmin asiakkaan heikko fyysinen toimintakyky ja heikko kognition taso. Tutkimusta varten teimme myös yhdistetyn luokittelun fyysisistä toimintakykyä ja kognition tasoa kuvaavista mittareista. Taulukossa 5 on esitetty hoitohenkilökunnan palkoilla painotetut hoitoajat RAI-järjestelmään perustuvien seitsemänluokkaisten fyysisistä toimintakykyä (ADL) ja kognitiota (CPS) mittavien mittarien luokissa. Korkeampi ADL-pistemäärä merkitsee suurempaa fyysisen avun tarvetta ja korkeampi CPS-pistemäärä enemmän heikentynyttä kognitiota. Hoitoajat kasvasivat selvästi fyysisen toimintakyvyn heikentyessä. Tämä pätee myös kognition osalta parhaimmassa ADL-luokissa (0–2), mutta ei enää ADL:n luokissa 3–6. Niistä

TAULUKKO 5. Hoitohenkilökunnan palkoilla painotettu hoitoaika minuuteissa vuorokauden aikana asiakkaiden fyysisen toimintakyvyn (ADL) ja kognition (CPS) mukaan¹

	ADL 0-2	ADL 3-4	ADL 5-6
CPS 0-2	68	104	115
CPS 3-4	77	97	107
CPS 5-6	87	94	106

¹ Suuri pistemäärä merkitsee heikkoa fyysistä tai kognitiivista toimintakykyä.

asiakkaista, jotka tarvitsivat runsaasti fyysistä apua tai jotka olivat täysin autettavia (ADL 3–6), kognitiiviselta kyvyltään terveemmät saivat hieman enemmän hoitoaika kuin vähintään lievistä kognition heikentymästä kärsivät asiakkaat.

Lisäksi masennusoireet ja päivittäinen kipu lisäsivät kumpikin hoitoaikaa noin 10 minuuttia vuorokaudessa ja hoitajien antama kuntoutus noin 5 minuuttia, muiden tekijöiden pysyessä muuttumattomina. Asiakkaan korkea ikä puolestaan vähensi hoitoaikaa. Jos asiakas oli 75–84-vuotias, väheni hoitoaika noin 12 minuuttia vuorokaudessa verrattuna alle 65-vuotiaisiin. Yli 85-vuotiailla hoitoaika väheni yli 15 minuuttia vuorokaudessa, kun muut asiakkaan ominaisuudet oli vakioitu. Tulos kertoo siitä, että laitoshoidossa nuorimmat asiakkaat tarvitsevat usein enemmän voimavaroja kuin kaikkein iäkkäimmät. Pitkään osastolla olleet asiakkaat saivat puolestaan hieman vähemmän hoitoaikaa kuin vähemmän aikaa osastolla hoidossa olleet asiakkaat. Akuuttiasiakkaat tarvitsevat siten keskimäärin enemmän voimavaroja kuin pitkäaikaisasiakkaat, mikä havaittiin myös osajulkaisussa 2. Lisäksi asiakkaan kanssakäynnin ongelmat muiden kanssa vähensivät hoitoaikaa noin 5 minuuttia vuorokautta kohden.

Osaston toimintaa ja rakennetta kuvaavat tekijät, kuten esimerkiksi osaston hoitopaikkamäärä, moduulityöskentelyn soveltaminen tai osaston erikoistuminen psykiatristen asiakkaiden ja dementia-asiakkaiden hoitoon lisäsivät hieman asiakkaiden hoitoaikaa, mutta tuloksiin liittyi tilastollista epävarmuutta. Osastotekijöistä ainoastaan hoitotyön osuus osastohenkilöstön kokonaistyöajasta lisäsi tilastollisesti merkitsevästi asiakkaan saamaa hoitoaikaa. Jos nykyistä hoitotyön osuutta kokonaistyöajasta, mikä on 49 prosenttia, pystyttäisiin nostamaan 59 prosenttiin, saisi kukin hoidossa oleva asiakas keskimäärin noin 7 minuuttia enemmän hoitoa vuorokaudessa. Tämä tarkoittaisi, että palkkapainotettu hoitoaika vuorokaudessa lisääntyisi keskimääräisestä 99 minuutista noin 106 minuuttiin asiakasta kohden.

Osastojen väliset erot työajan kohdentumisessa säilyivät suurina vaikka niiden erilainen asiakasrakenne otettiin huomioon. Eroja ei kuitenkaan pystytty juurikaan selittämään osastojen erilaisen rakenteen ja toimintaympäristön avulla. Laskimme myös keskimääräisen asiakaskohtaisen hoitoajan ennusteen kullekin osastolle osaston asiakasrakenteen perusteella ja vertasimme sitä toteutuneeseen keskimääräi-

seen asiakkaan hoitoaikaan osastossa. Ero ennustetun ja havaitun hoitoajan välillä vaihteli vuorokaudessa -30 minuutista +30 minuuttiin, mikä kertoo siitä, että ei-mitattavissa olevat asiakastason tekijät, työpanoksen erilainen käyttö ja satunnaisvaihtelu vaikuttavat hoitoaikoihin ja sitä kautta osaston välisiin eroihin. Ennustetun ja toteutuneen hoitoaikojen erotuksen jakauma muistuttikin hajonnaltaan osaston tuotannollisen tehokkuuden jakaumia. Monitasomallissa vakioidut asiakas- ja osastokohtaiset tekijät selittivät hieman alle 30 prosenttia hoitoajan vaihtelusta, mikä oli suhteellisen vähän verrattuna aiempiin tutkimuksiin (Björkgren et al., 1999). Asiakkaan RUG-III/22-indeksi selitti yksittäisiä MDS 2.0-lomakkeesta saatavia asiakkaan ominaisuuksia kuvaavia tietoja paremmin hoitoaika. Tämä puoltaa sitä, että asiakasrakennetta kuvaavana mittarina tulisi käyttää mittaria, joka kuvaa mahdollisimman laaja-alaisesti asiakkaan ominaisuuksia. Pelkkä fyysisen toimintakyvyn mittaaminen ei ole riittävää, varsinkin jos halutaan arvioida asiakaiden ominaisuuksien yhteyttä voimavaratarpeeseen tai asiakaskohtaisiin kustannuksiin.

11 Pohdinta

11.1 Laitoshoidon tuotannollinen tehokkuus ja hoidon laatu

Tuotannollisen tehokkuuden, hoidon laadun ja asiakkaille annettavan hoidon määrän erot eri osastojen välillä olivat erittäin suuria ja merkittäviä. Tulokset osoittivat myös, että tuotannollisen tehokkuuden ja hoidon laadun välillä on vaihtosuhte (trade-off). Tuotannollinen tehokkuus saatetaan saavuttaa hoidon laadun kustannuksella, mikä olisi huomioitava tuottavuusvertailuissa ja sekä toiminnan strategisessa että operatiivisessa johtamisessa. Tuotannollista tehokkuutta mitattiin tässä tutkimuksessa teknisenä tehokkuutena ja kustannustehokkuutena ja niihin vaikuttavat merkittävästi osaston henkilöstömäärä ja -mitoitus. Tutkimustulokset kertovat näin myös siitä, että yksikön työvoiman käyttö ja toiminnan resursointi tai ainakin henkilöstövoimavarojen optimaalisessa käytössä epäonnistuminen ovat hoidon laadun ohella osastoja voimakkaasti erottelevia tekijöitä.

Laadun yhteys tuotannolliseen tehokkuuteen ei kuitenkaan keskimäärin ole niin voimakas kuin usein ajatellaan. Laadun ja tuotannollisen tehokkuuden välinen yhteys on havaittavampi ja merkittävämpi silloin, kun näiden tekijöiden suhteessa tapahtuu suuria muutoksia tai jos yksikkö on muihin verrattuna poikkeuksellisen tehokas tai laadukas. Pitkäaikaishoitoyksiköiden näkökulmasta tulos on sikäli myönteinen, että pienet muutokset esimerkiksi hoidon kliinisessä laadussa eivät välttämättä näy merkittävästi tuotannollisessa tehokkuudessa tai kustannuksissa. Toisaalta myös tuotannollisessa tehokkuudessa voi tapahtua kehitystä ilman, että hoidon laadussa tapahtuu haitallisesti merkittäviä muutoksia. Näin ollen on myös mahdollista, että eräät tutkimuksessa mukana olleista tehottomista yksiköistä voivat jossakin määrin parantaa tuottavuuttaan ilman, että hoidon laadussa tapahtuu merkittäviä negatiivisia muutoksia ja päinvastoin. Ns. benchmarking-toiminnassa (vertailukehittämisessä) ja suorituskyvyn arvioinnissa tulee kuitenkin kiinnittää erityistä huomiota sellaisiin yksiköihin, jotka ovat muihin nähden joko tuotannollisen tehokkuuden tai laadun suhteen poikkeavia.

11.2 Tuotannollisen tehokkuuden erojen kaventaminen ja tuottavuuden parantaminen

Vanhusten pitkäaikaisen laitoshoidon kannalta on tärkeää, että itse palvelujen tuotanto on tehokasta (tuottavuus) ja että palvelut ovat laadukkaita ja vaikuttavia. Tämän tutkimuksen mukaan osastojen voimavaroja ei kaikissa yksiköissä ole käytetty optimaalisesti suhteessa tuotettuihin hoitopäiviin eikä henkilöstömitoititus ole suhteessa asiakasrakenteeseen. Osastojen tuottavuudessa on suuria eroja, vaikka osastojen erilainen asiakasrakenne otetaan huomioon. Vanhusten pitkäaikaishoidossa onkin kiinnitettävä huomiota keinoihin, joiden avulla työvoiman käyttöä ja sen oikeudenmukaista jakautumista eri osastoille voidaan parantaa. Tällä hetkellä osastojen väliset erot kertovat jossakin määrin epätarkoituksenmukaisesta voimavarojen käytöstä ja voimavarojen kohdentamisen ja optimoinnin ongelmista.

Tuottavuutta voidaan parantaa makrotasolla muuttamalla palvelurakennetta eli esimerkiksi lisäämällä kunnassa kotihoidon ja palveluasumisen osuutta suhteessa kalliimpaan vanhainkoti- ja terveyskeskushoitoon tai edistämällä ikääntyneen väestön omatoimisuutta (Vaarama & Voutilainen, 2002). Tällöin kyse on kunnan palvelutuotantostrategiaan ja palvelujen organisointiin liittyvistä muutoksista. Tämä tutkimus kuitenkin osoitti selkeästi, että tuotannollisen tehokkuuden erot saman toimialan yksiköiden välillä ovat merkittäviä. Vanhustenhuollon tuottavuuteen ja vanhuspalvelumenoihin vaikuttaa siten merkittävästi myös yksittäisten yksikköjen toiminta. Yksittäisen yksikön näkökulmasta kyse on tällöin ennen kaikkea tuotantoprosessiin liittyvistä tekijöistä. Tuottavuuden parantaminen ei voisi-kaan tapahtua vain makrotason toimenpiteiden avulla, vaan muutoksia tarvittaisiin myös mikrotasolla. Nämä näkökulmat yhdistyvät luontevasti, jos vanhusten palveluja arvioidaan hoitoketjuina, joissa tavoitteena on hoitaa asiakkaat kustannus-vaikuttavasti riippumatta siitä, missä palvelurakenne- ja yksikkötasolla (vastuutasolla) asiakas kulloinkin on hoidossa.

Tuottavuuden parantaminen on siten myös yksittäisten yksiköiden vastuulla ja vaikutettavissa oleva asia. Hoidon laatu, hoitopaikkojen kuormitusaste, tuotettujen hoitopäivien määrä, työvoiman hinta ja määrä, osastotyyppi ja työajan käyttö vaikuttavat osastojen tuotannolliseen tehokkuuteen. Osa näistä tekijöistä saattaa olla yksittäisen osaston vaikutusmahdollisuuksien ulkopuolella, mutta jossakin määrin niihin varmasti pystytään vaikuttamaan osasto- ja laitostasolla. Yksittäisen yksikön tehokkuuteen vaikuttavat myös palkitsemis- ja rahoitusjärjestelmään liittyvät kannustimet, joiden osalta vanhustenhuollossa on paljon kehittämismahdollisuuksia. Julkisten yksiköiden tavoitteet ovat moninaiset ja ainakin tuotannollisen tehokkuuden ja hoidon laadun välillä on vaihtosuhte. Tästä syystä tarvitaan kannuste- ja seurantajärjestelmiä, joilla sekä yksittäisen yksikön toimintaa että koko pitkäaikaishoitoa voidaan arvioida ja ohjata haluttuun suuntaan. Suomessa on

käytössä erinomaisia välineitä asiakasrakenteen, tuottavuuden ja hoidon laadun systemaattiseen arviointiin ja niitä voitaisiin hyödyntää selvästi nykyistä enemmän sekä kunta- että yksiköitasolla.

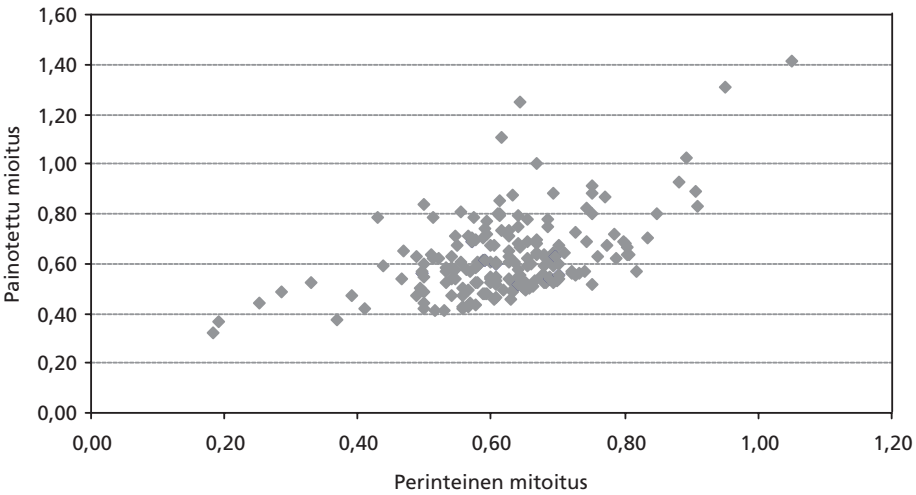
Tuottavuuden parantaminen myös vanhusten pitkäaikaisessa laitoshoidossa on ilmeisesti terveystalouden keskeinen tavoite tulevaisuudessa (VM, 2003). On kuitenkin muistettava, että tuottavuutta voidaan parantaa ilman, että kaikkein tehokkaimmat yksiköt tehostavat toimintaansa entistäkin enemmän. Tuottavuuden määritelmän mukaisesti tuottavuus saadaan nousemaan sillä, että tuotannollisesti kaikkein tehottomimmat yksiköt pystyvät parantamaan tuotannollista tehokkuuttaan. Joka tapauksessa tehokkuuspyrkimykset edellyttävät laadun seuranta- ja arviointijärjestelmien käyttöä, jotta parempaan tuottavuuteen tähtäävien muutosten seurauksia voidaan arvioida laajasti sekä yhteiskunnan, tuottajien ja asiakkaiden näkökulmista. Tuottavuuden parantaminen ei ole hyväksyttävää, jos se tapahtuu hoidon laadun tai vaikuttavuuden kustannuksella. Tehokkuuserojen kaventaminen ja tuottavuuden parantaminen edellyttävät myös sitä, että yksiköt tuntevat tuotantoprosessinsa ja siinä vaikuttavat olennaisimmat tekijät ja näkevät tuottavuuden tärkeänä asiana. Vaikka vanhusten pitkäaikaisessa laitoshoidossa näyttääkin olevan runsaasti kehittämismahdollisuuksia, tilanne ei välttämättä ole sen huonompi kuin muillakaan sosiaali- ja terveydenhuollon toimialoilla. Esimerkiksi tuotannollisen tehokkuuden erot eri sairaaloiden tai eri terveyskeskusten välillä on lähes vastaavanlaiset kuin vanhustenhuollossa (Hjerpe et al., 2003). Lisäksi muissakin terveyspalveluissa tuottavuuden kehitys on viime vuosina ollut laskeva (Hjerpe et al., 2003). Vanhustenhuollon laatuakaan ei Suomessa voida kokonaisuudessaan pitää huonona. Ainakin Yhdysvaltoihin verrattuna vanhusten pitkäaikaishoidon laatu on Suomessa vähintään kelvollinen (Finne-Soveri, 2005). Vanhustenhuollon palvelujen järjestämiseksi tarvitaan laitoshoidon lisäksi vertailukelpoista tietoa myös kotihoidosta eli kotipalveluista ja kotisairaanhoidosta. Vanhustenhuollon laaja-alainen kehittäminen edellyttää kunnassa koko palvelurakenteen ja hoitoketjujen tarkastelua. Laitoshoido on vain yksi osa laajaa kokonaisuutta.

11.3 Henkilöstön sijoittelun ja työajan käytön kehittäminen

Henkilöstövoimavarojen oikeudenmukainen jakaminen laitoksen sisällä, hoidon kohdistaminen eniten hoitoa tarvitseville asiakkaille ja osaston henkilöstön palkitseminen hyvästä laadusta ovat keskeisiä keinoja, joilla vanhusten laitoshoidoa voidaan tulevaisuudessa kehittää. Pelkästään henkilöstömitoituksen lisääminen kaikissa yksiköissä ei välttämättä johda laadullisesti tai määrällisesti parempaan

hoitoon, jos palvelutuotantoa ei tarkastella kokonaisuutena. Tuloksena voi olla, että tuottavuuden lasku ylittää saadut hyvinvointihyödyt. Alhaista henkilöstömitoitusta pidetään kuitenkin laitoshoidon suurimpana ongelmana.⁹

Tehokkuuslukujen laskemiseksi käytettiin tässä tutkimuksessa tuotannontekijöinä henkilöstömääriä tai henkilötyövuosia. Tuotoksena käytettiin hoitopäiviä, joita painotettiin asiakasrakenteella. Näin ollen yksiköiden tuotannollisen tehokkuuden eroihin vaikuttaa myös se, miten henkilöstö ja asiakkaat on osastoille sijoitettu. Perinteistä henkilöstömitoituksen laskentatapaa voitaisiin laajentaa oikeudenmukaisempaan, osastojen voimavaratarvetta vastaavampaan ja tuotannollista tehokkuutta edistävämpään suuntaan siten, että laskennassa huomioitaisiin selkeästi osaston asiakasrakenne. Kuviossa 13 on esitetty perinteisen henkilöstömitoituksen ja asiakasrakenteella painotetun mitoituksen yhteys 195 osastossa vuonna 2002. Perinteinen henkilöstömitoitus on laskettu jakamalla osaston henkilöstömäärä osaston hoitopaikkamäärällä. Painotettu mitoitus on laskettu samalla tavalla, mutta osaston hoitopaikkamäärää on painotettu osaston asiakkaiden keskimääräisellä RUG-III/22-indeksillä. Painotettu mitoitus on perinteistä pienempi, jos osaston asiakasrakenne on keskimääräistä indeksii 1,0 suurempi. Esimerkiksi jos osaston henkilöstömäärä on 10, hoitopaikkoja on 25 ja osaston asiakkaiden keskimääräinen kustannuspaino 1,25, painotettu mitoitus on $10 / (25 * 1,25) = 0,32$. Perinteinen tai keskimääräisellä indeksillä 1,0 painotettu mitoitus osastolla olisi



KUVIO 13. Perinteinen ja asiakasrakenteella painotettu henkilöstömitoitus

⁹ Vuoden 2001 lopulla RAI-tietojärjestelmän käyttöönotto ja pitkäaikaishoidon benchmarking-hankkeessa toteutetun kyselyn perusteella 75 prosenttia osastonhoitajista piti osastonsa sen hetkistä henkilöstömitoitusta liian pieninä. Myös vanhustenhoidon laatusuosituksissa henkilöstömitoituksen arvioinnilla näyttää olevan merkittävä painoarvo toiminnan laadun arvioinnissa (Vaarama et al., 2001).

puolestaan $10 / 25 = 0,4$. Kuviossa 13 esitetystä sirontakuviosta korrelaatio mitoitustulukujen välillä on vain 0,52. Kuviosta havaitaankin, että perinteinen mitoitustulos ei noudata kaikilta osin asiakasrakenteella painotettua mitoitusta, joten henkilöstön ja asiakkaiden sijoittelussa osastoille on kehittämismahdollisuuksia. Molempien lukujen keskiarvo on 0,62, mutta painotetun mitoitustuloksen keskihajonta on perinteistä mitoitustulosta suurempi, mikä sekin kertoo henkilöstön uudelleenalkamisen tarpeesta laitoksien eri osastojen välillä. Henkilöstön jakautumiseen vaikuttavat myös muut tekijät kuin asiakasrakenne kuten esimerkiksi osastojen fyysiset puitteet, jotka tulee luonnollisesti ottaa paikallisesti huomioon.

11.4 Tuotannollinen tehokkuus ja etiikka

Tämän tutkimuksen tavoitteena oli soveltaa perinteistä taloustieteellistä tuottavuustutkimusta vanhusten laitoshoidon tutkimukseen ja samalla huomioida aiempia tutkimuksia systemaattisemmin ja kattavammin myös hoidon laatu. Tuotannollinen tehokkuus ja hoidon laatu saatetaan kokea ristiriitaisina tavoitteina ja palvelutuotannon tehokkuuteen ja tuottavuuteen liittyvät pyrkimykset jopa epäeettisinä. Tuotannollista tehokkuutta ja sen tavoittelemista ei kuitenkaan lähtökohdaisesti tulisi pitää arveluttavana, vaan pikemminkin välttämättömänä tavoitteena voimavarojen tehokkaan käytön saavuttamiseksi. Kysymys onkin pikemminkin siitä, että päätöksenteossa tulisi tiedostaa se, millainen yhteys tuotannollisella tehokkuudella on hoidon laatuun ja vaikuttavuuteen ja miten muutokset näissä tekijöissä vaikuttavat toisiinsa. Suomessa on kansainvälisesti erittäin hyvät mahdollisuudet arvioida vanhusten laitoshoidon niin hoidon laadun, vaikuttavuuden kuin palvelutuotannon tehokkuudenkin näkökulmasta. Tämä mahdollisuus tulisi käyttää hyödyksi päätöksenteossa ja käytännön hoitotyössä. Laatu ja tuottavuus ovat kaiken palvelutuotannon olennaisia osa-alueita. Niitä on tarkasteltava yhtäaikaaisesti, avoimesti ja rohkeasti myös erilaisissa hoivapalveluissa.

Tämän tutkimuksen tuotannollista tehokkuutta koskevat tulokset ja johtopäätökset tulee nähdä myös oikeudenmukaisuuden valossa. Osastojen väliset erot hoidon laadussa, tuotannollisessa tehokkuudessa ja asiakkaille kohdennetun hoidon määrässä kertovat voimavarojen kohdentamiseen ja käyttöön liittyvistä epäkohdista. Yksiköiden väliset erot ovat merkittävän suuria. Tulokset kertovat myös siitä, että eräät osastot ilmeisesti joutuvat toimimaan selvästi toisia osastoja tehokkaammin eli lähempänä tehokkuusrintamaa. Osaston tuotannollinen tehokkuus voi jossakin määrin olla ulkoapäin annettu, ainakin silloin jos osaston voimavarat, asiakasmäärä ja asiakasrakenne ovat eksogeenisia. Syistä riippumatta osastojen väliset erot ovat joka tapauksessa sekä hoidossa olevien asiakkaiden, hoitotyöntekijöiden, rahoittajien ja koko palvelujärjestelmän toimivuuden kannalta ongelmallisia.

11.5 Tutkimukseen liittyvät varaukset ja jatkotutkimustarpeet

Tuotoksen mittaamista ja moninaisuutta on erityisesti terveydenhuollossa pidetty eräänä keskeisimpänä tuottavuustutkimuksiin liittyvänä ongelmana (Hjerppe & Luoma, 2003; Newhouse, 1994). Tuotoksen moninaisuuden aiheuttamaa ongelmaa pyrittiin tässä tutkimuksessa minimoimaan vakioimalla tuotoksessa yksiköiden asiakasrakenne. Pitkäaikaishoidossa tuotos on toki jo lähtökohtaisestikin homogeenisempi kuin esimerkiksi erikoissairaanhoidossa. Tehokkuusluvut laskettiin poikkileikkausaineistosta, jossa kunkin osaston tuotosta eli hoitopäiviä painotettiin osaston asiakkaiden keskimääräisellä RUG-III/22-indeksillä. Tuotoksen vakiointi tällä tavoin on vielä toistaiseksi harvoin sovellettu mutta erinomainen keino saattaa eri osastojen tuotos vertailukelpoiseksi. Tuotosta ei pyritty painottamaan laadulla, koska käytössä ei ollut riittävän kattavaa yksittäistä laatumittaria, jonka avulla hoitopäiviä olisi voitu painottaa.

Tutkimuksessa käytettiin tuotantoteoriaan perustuvan tehokkuustutkimuksen vakiintuneita menetelmiä, joihin kuitenkin liittyy eräitä tunnettuja ongelmia. Stokastiset rintamamallit ovat herkkiä parametreihin liittyville oletuksille ja DEA-menetelmä puolestaan aineisto- ja mittausvirheille (Dyson et al., 2001; Newhouse, 1994). Yksittäisten yksiköiden tehokkuuslukuihin liittyy myös jonkin verran epävarmuutta poikkileikkausaineiston vuoksi. Tehokkuuslukuihin liittyvää epävarmuutta voidaan vähentää käyttämällä paneeliaineistoa (Street, 2003), jota ei kuitenkaan ollut kattavasti saatavissa tähän tutkimukseen. Aivan viimeaikaiset tutkimukset ovat kuitenkin osoittaneet, että yksiköiden suhteellinen asema tehokkuusvertailuissa säilyisi vuodesta toiseen suhteellisin muuttumattomana (Laine, 2005b). Eri vuosien poikkileikkausaineistot antavat yhdenmukaisen ja luotettavan kuvan toimialan tuotannollisesta tehottomuudesta 2000-luvulla. Tulevaisuudessa vanhusten laitoshoidon arvioinnissa tulee kuitenkin kiinnittää enemmän huomiota paneeliaineistojen käyttöön ja hoidon kustannus-vaikuttavuuteen, jonka avulla voidaan yhdistää aiempaan paremmin hoidon laadulliset, ajalliset ja taloudelliset tekijät. Vasta tällöin voidaan kattavasti arvioida sitä, missä yksiköissä asiat ovat loppujen lopuksi hyvin ja missä huonosti.

Tutkimusaineisto sisälsi noin 120 osastoa, joissa oli hoidossa keskimäärin noin 3600 asiakasta. Vaikka tämä on vain noin 11 prosenttia kaikista vanhainkodeissa ja terveyskeskussairaaloiden vuodeosastoissa hoidettavista asiakkaista, voidaan tulokset yleistää koskemaan koko maata. RAI-tietojärjestelmän käyttöönotto ja pitkäaikaishoidon benchmarking-hankkeessa saatujen kokemusten perusteella hankkeessa mukana olevien osastojen määrän vuosittainen lisääntyminen ei ole vaikuttanut merkittävästi keskeisimpien tunnuslukujen (esim. laatu, tuotannollisen tehokkuus) keski- ja hajontalukuihin (Finne-Soveri, 2005; Laine, 2005b). On

kuitenkin muistettava, että tehokkuuslukujen laskeminen perustui empiiriseen aineistoon, joten tehokkaimpien yksiköiden muodostama tehokkuusrintama ei ole teoreettinen, vaan empiirinen konstruktio. Näin ollen on mahdollista, että tutkimuksessa ei ollut mukana kaikista Suomen pitkäaikaishoito-osastoista tehokkaimpia yksiköitä, jolloin mikään tutkimuksessa ollut osasto ei siten toimisi todellisella tehokkaan tuotannon rintamalla. Tässä tutkimuksessa tehokkuusrintamalla olevat yksiköt saattavatkin tosiasiaassa pystyä parantamaan sekä hoidon laatuaan että tuotannollista tehokkuuttaan nykyisellä tuotantoteknologiallaan ja voimavaroillaan.

Knappin ym. (Knapp, 1984; Knapp & Davies, 1981) mukaan laitoshoidon taloudellisessa tarkastelussa tulisi ottaa huomioon myös aineettomat voimavarat, joita ovat esimerkiksi hoitomiljöö, työympäristö, henkilöstö ja organisaatorakenteet. On mahdollista, että osa havaituista tuotannollisen tehokkuuden ja työajan käytön eroista voidaan selittää esimerkiksi henkilöstön hyvinvointiin liittyvillä tekijöillä. Työyhteisön hyvinvointiin liittyviä tekijöitä ei kuitenkaan ollut mahdollista ottaa mukaan tähän tutkimukseen. Pekkarisen ym. tutkimuksessa (Pekkarinen et al., 2004) havaittiin, että psykiatristen asiakkaiden ja dementia-asiakkaiden hoitoon erikoistuneiden osastojen henkilöstön hyvinvointi oli parempi kuin sekaosastojen henkilöstön hyvinvointi. Tässä tutkimuksessa havaittiin, että sekaosastot toimivat kuitenkin erikoistuneita osastoja tehokkaammin (tekniinen tehokkuus). Osaston erikoituminen saattaakin lisätä henkilöstön hyvinvointia mutta vähentää teknistä tehokkuutta. Toiminnan johtamisen näkökulmasta tällainen yhteys on luonnollisesti haastava. Tuottavuuden heikkeneminen voi siten olla seuraus toiminnan muunlaisesta kehittämisestä. Tulevaisuudessa laitoshoidon arvioinnissa olisikin pyrittävä huomiomaan myös aineettomat voimavarat aiempaa laajemmin. Henkilöstön hyvinvointi on itse asiassa laatuakin vähemmän huomioitu tekijä perinteisissä tuottavuustutkimuksissa.

Yksityisten sosiaali- ja terveyspalvelujen määrän lisääntyminen Suomessa on ollut suhteellisen voimakasta viime vuosina. Vanhustenhuollossa yksityiset yritykset ovat keskittyneet tarjoamaan pääasiassa kotipalveluja ja palveluasumista. Tämän tutkimuksen perusteella ei voida ottaa kantaa siihen ajankohtaiseen kysymykseen tai olettamukseen, pystyisikö markkinaperusteinen palvelutuotanto tuottamaan palveluita julkista tuotantoa kustannus-vaikuttavammin tai tuotannollisesti tehokkaammin? Runsaasti voimavaroja vaativien vanhuspalvelujen tuottaminen (esimerkiksi terveyskeskushoito) on ja tulee varmasti olemaankin edelleen suurelta osin julkisen sektorin vastuulla. Tähän tutkimukseen ei ollut mahdollista saada riittävää määrää yksityistä tuottajia, jotta tätä kysymystä olisi voitu tutkia. Eräät kansainväliset tutkimukset ovat antaneet viitteitä siitä, että yksityinen palvelutuotanto ja todellisen kilpailun olemassaolo edistäisivät (kustannus)tehokkuutta, mutta toistaiseksi vastaavasta ei ole näyttöä Suomen pitkäaikaishoidossa.

LÄHTEET

- Aaltonen, J., Järviö, M.-L., Luoma, K. & Rätty, T. (2004). Terveyskeskusten tuottavuuden ja tehokkuusrojen kehitys vuosina 1988–2002. VATT, Keskustelunaloitteita 354. Helsinki.
- Aigner, D. J., Lovell, C. A. K. & Schmidt, P. (1977). Formulation and Estimation of Stochastic Frontier Production Function Models. *Journal of Econometrics* 6(6): 21–37.
- Alasuutari, P. (2004). Suunnitelmataloudesta kilpailutalouteen: miten muutos oli ideologisesti mahdollinen? *Yhteiskuntapolitiikka* 69(1): 3–16.
- Anderson, R. I., Lewis, D. & Webb, J. R. (1999). The Efficiency of Nursing Home Chains and the Implications of Non-profit Status. *Journal of Real Estate Portfolio Management* 5(3): 235–245.
- Arrow, K. (1963). Uncertainty and the welfare economics of medical care. *American Economic Review* 53(5): 941–973.
- Barr, N. (1998). *The Economics of the Welfare State*. Third Edition ed. Oxford University Press: Oxford.
- Battese, G. & Broca, S. (1997). Functional Forms of Stochastic Frontier Production Functions and Models for Technical Inefficiency Effects: A Comparative Study for Wheat Farmers in Pakistan. *Journal of Productivity Analysis* 8(4): 395–414.
- Battese, G. & Coelli, T. (1993). *A Stochastic Frontier Production Function Incorporating a Model for Technical Inefficiency Effects*. University of New England. Armidale, Australia, 69.
- Battese, G. & Coelli, T. (1995). A Model for Technical Inefficiency Effects in Stochastic Frontier Production Function for Panel Data. *Empirical Economics* 20: 325–332.
- Baumol, W. (1967). Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crises. *American Economic Review* 57(3): 415–426.
- Berg, K., Mor, V., Morris, J., Murphy, K., Moore, T. & Harris, Y. (2002). Identification and Evaluation of Existing Nursing Homes Quality Indicators. *Health Care Financing Review* 23(4): 19–36.
- Björkgren, M. (2002). Case-mix Classification and Efficiency Measurement in Long-Term Care of the Elderly. STAKES, Research report 124. Helsinki.
- Björkgren, M., Häkkinen, U. & Finne-Soveri, H. (1998). Pitkäaikaishoidon voimavaratarve RUG-luokituksella. Stakes, Aiheita 1. Helsinki.
- Björkgren, M., Häkkinen, U., Finne-Soveri, H. & Fries, B. (1999). Validity and reliability of Resource Utilization Groups (RUG-III) in Finnish long-term care facilities. *Scandinavian Journal of Public Health* 27(3): 228–234.
- Björkgren, M., Häkkinen, U. & Linna, M. (2000). A comparison of methods for determining efficiency of Finnish long-term care units. *Sosiaalilääketieteellinen aikakauslehti* 37: 216–222.
- Björkgren, M., Häkkinen, U. & Linna, M. (2001). Measuring Efficiency of Long-Term Care Units in Finland. *Health Care Management Science* 4(3): 193–200.
- Blank, J. L. T. & Eggink, E. (2001). A Quality-Adjusted Cost Function in a Regulated Industry: The Case of Dutch Nursing Homes. *Health Care Management Science* 4(3): 201–211.
- Carpenter, I., Perry, M., Challis, D. & Hope, K. (2003). Identification of registered nursing care of residents in English nursing homes using the Minimum Data Set Resident Assessment Instrument (MDS/RAI) and Resource Utilisation Groups version III (RUG-III). *Age and Ageing* 32(3): 279–285.
- Carrillo, E., Garcia-Altes, A., Peiro, S. & Portella, E. (1996). System for the classification of patients in mid and long-term care facilities: Resource Utilization Groups, version III. Validation in Spain. *Revista de Gerontologia* 6(4): 276–284.
- Charnes, A., Cooper, W. W. & Rhodes, E. (1978). Measuring the Efficiency of Decision Making Units. *European Journal of Operational Research* 2: 429–444.
- Chattopadhyay, S. & Ray, S. C. (1996). Technical, scale, and size efficiency in nursing home care: a non-parametric analysis of Connecticut homes. *Health Economics* 5(4): 363–373.
- Ching-To, A. (1994). Health Care Payment Systems: Cost and Quality Incentives. *Journal of Economics & Management Strategy* 3(1): 93–112.
- Chou, S. Y. (2002). Asymmetric information, ownership and quality of care: an empirical analysis of nursing homes. *Journal of Health Economics* 21(2): 293–311.
- Coelli, T. (1996). *A Guide to FRONTIER version 4.1: A computer program for stochastic frontier production and cost function estimation*. University of New England. Armidale, Australia, CEPA Working Paper 7.

- Coelli, T., Rao, D. & Battese, G. (1998). *An Introduction to Efficiency and Productivity Analysis*. Fourth ed. Kluwer Academic Publishers.
- Cooper, W. W., Seiford, L. M. & Tone, K. (2000). *Data Envelopment Analysis. A Comprehensive Text with Models, Applications, References and DEA-Solver Software*. Kluwer Academic Publishers.
- Crivelli, L., Filippini, M. & Lunati, D. (2002). Regulation, Ownership and Efficiency in the Swiss Nursing Home Industry. *International Journal of Health Care Finance and Economics* 2(2): 79–97.
- Croxson, B., Propper, C. & Perkins, A. (2001). Do doctors respond to financial incentives? UK family doctors and the GP fundholder scheme. *Journal of Public Economics* 79(2): 375–398.
- Davis, K. (1972). Economic Theories of Behavior in Nonprofit, Private Hospitals. *Economic and Business Bulletin* 24(2): 1–13.
- Dawson, D., Gravelle, H., Kind, P., O'Mahony, M., Street, A. & Weale, M. (2004). *Developing New Approaches to Measuring NHS Outputs and Productivity*. Centre for Health Economics. The University of York, CHE Technical Paper Series 31. York.
- Dellefield, M. E. (2000). The Relationship Between Nurse Staffing in Nursing Homes and Quality Indicators. *Journal of Gerontological Nursing* 27(6): 14–28.
- Donabedian, A. (1988). The Quality of Care? How Can It Be Assessed? *JAMA* 260(12): 1743–1748.
- Drummond, M. F., O'Brien, B., Stoddart, G. L. & Tarrance, G. W. (1997). *Methods for the Economic Evaluation of Health Care Programmes*. Second ed. Oxford University Press Inc.: New York.
- Dyson, R. G., Allen, R., Camanho, A. S., Podinovski, V. V., Sarrico, C. S. & Shale, E. A. (2001). Pitfalls and protocols in DEA. *European Journal of Operational Research* 132(2): 245–259.
- Eggleston, K., Grossman, J. & Cutler, D. (2004). Productivity Research and Healthcare Delivery Innovation. *Applied Health Economics and Health Policy* 3(3): 133–141.
- Evans, R. & Stoddart, G. L. (1990). Producing health, consuming health care. *Social Science & Medicine* 31(12): 1347–1363.
- Farrell, M. J. (1957). The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society A* 120(3): 253–281.
- Finne-Soveri, H. (1999). Pitkääikaishoidon laatu on mitattavissa, vertailtavissa ja parannettavissa. *Duodecim* 115: 1600–1610.
- Finne-Soveri, H. (2005). Hoidon laatu pitkääikaaisessa laitoshoidossa. Teoksessa A. Noro, H. Finne-Soveri, M. Björkgren & P. Vähäkangas (toim.). *Ikääntyneiden laitoshoidon laatu ja tuottavuus - RAI-järjestelmä vertailukehittämässä*. Stakes, M 205. Helsinki.
- Fizel, J. L. & Nunnikhoven, T. S. (1992). Technical efficiency of for-profit and non-profit nursing homes. *Managerial and Decision Economics* 13(5): 429–439.
- Folland, S. T. & Hofer, R. A. (2001). How reliable are hospital efficiency estimates? Exploiting the dual to homothetic production. *Health Economics* 10(8): 683–698.
- Fries, B. E., Cooney, J. R. & Leo, M. (1985). Resource Utilization Groups. A Patient Classification System for Long-term Care. *Medical Care* 23(2): 110–122.
- Fries, B. E., Schneider, D. P., Foley, W. J., Gavazzi, M., Burke, R. & Cornelius, E. (1995). Refining a Case-Mix Measure for Nursing Homes: Resource Utilization Groups (RUG-III). *Medical Care* 32(7): 668–685.
- GAO. (2002). *Nursing Homes - Quality of Care More Related to Staffing than Spending*. The General Accounting Office, GAO-02-431R. Washington.
- Gaynor, M. (1994). Issues in the Industrial Organization of the Market for Physician Services. *Journal of Economics & Management Strategy* 3(1): 211–249.
- Gertler, P. J. & Waldman, D. M. (1992). Quality-adjusted Cost Functions and Policy Evaluation in the Nursing Home Industry. *Journal of Political Economy* 100(6): 1232–1256.
- Goldstein, H. (1995). *Multilevel Statistical Models*. Second Edition ed. Edward Arnold: London.
- Greene, W. H. (1993). *Econometric Analysis*. Prentice-Hall, Inc.: New Jersey.
- Grossman, M. (1972). On the concept of capital and the demand for health. *Journal of Political Economy* 80(2): 823–255.
- Hansmann, H. (1987). Economic theories on nonprofit organization. In W. Powell (Ed.). *The nonprofit sector. A Research Handbook* Yale University Press. New Haven.
- Harrington, C., Zimmermann, D., Karon, S. L., Robinson, J. & Beutel, P. (2000). Nursing home staffing and its relationship to deficiencies. *Journal of Gerontology: Social Sciences* 55(5): 278–287.

- Hawes, C., Morris, J., Phillips, C., Mor, V., Fries, B. & Nonemaker, S. (1995). Reliability Estimates for The Minimum Data Set for Nursing Home Resident Assessment and Care Screening. *The Gerontologist* 35(2): 172–178.
- Heikkilä, M. & Uusitalo, H. (1997). *The Cost of Cuts. Stakes: Helsinki.*
- Hjerppe, R., Kangasharju, A. & Vuorento, R. (2003). Kunnalliset palvelut – Terveysten- ja vanhustenhuollon tuottavuus. VATT-julkaisuja 37. VATT: Helsinki.
- Hjerppe, R. & Luoma, K. (2003). Tuottavuuden mittaaminen – käsitteet. Teoksessa R. Hjerppe, A. Kangasharju & R. Vuorento (toim.). Kunnalliset palvelut – Terveysten- ja vanhustenhuollon tuottavuus. VATT, VATT-julkaisuja 37. Helsinki.
- Ikegami, N., Fries, B., Takagi, U., Ikeda, S. & Ibe, T. (1994). Applying RUG-III in Japanese long-term care facilities. *The Gerontologist* 34(5): 628–639.
- Jacobs, R. (2000). Alternative Methods to Examine Hospital Efficiency: Data Envelopment Analysis and Stochastic Frontier Analysis. Centre for Health Economics, Discussion paper 177. York, England.
- Kane, R. (1998). Assuring Quality in Nursing Home Care. *Journal of American Geriatrics Society* 46: 232–237.
- Karon, S. & Zimmermann, D. (1996). Using Indicators To Structure Quality Improvement Initiatives in Long-Term Care. *Quality Management in Health Care* 4(3): 54–66.
- Kauppinen, S. (2001). Yksityiset sosiaalipalvelut 2000 – tiedonantajapalaute 11/2001. Stakes, Helsinki.
- Knapp, M. (1984). *The Economics of Social Care.* Macmillan Education.: Hong Kong.
- Knapp, M. & Davies, B. (1981). *Old People's Homes and the Production of Welfare.* Routledge & Kegan Paul Ltd.: England.
- Knox, K. J., Blankmeyer, E. C. & Stutzman, J. R. (2003). Organizational Efficiency and Quality in Texas Nursing Facilities. *Health Care Management Science* 6(3): 175–188.
- Koopmans, T. (1951). *Activity analysis of production and allocation.* Wiley: New York.
- Kooreman, P. (1994). Nursing Home Care in The Netherlands: A Nonparametric Efficiency Analysis. *Journal of Health Economics* 13(3): 301–216.
- Kumbhakar, S., Ghosh, S. & McGuckin, J. (1991). A Generalized Production Frontier Approach for Estimating Determinants of Inefficiency in U.S. Dairy Farms. *Journal of Business & Economic Statistics* 9(3): 279–286.
- Lahtinen, Y. (2003). Hoidon tarve ja tuottavuus vanhustenhuollossa vuosina 1998–2001. Teoksessa R. Hjerppe, A. Kangasharju & R. Vuorento (toim.). Kunnalliset palvelut – Terveysten- ja vanhustenhuollon tuottavuus. VATT, VATT-julkaisuja 37. Helsinki.
- Laine, J. (1999). Selvitys sosiaalihuollon erityispalveluiden tuotantomalleista ja sopimusmenettelyistä. Stakes, Aiheita 38. Helsinki.
- Laine, J. (2003). Pitkäaikaishoidon tehokkuus ja laatu vuosina 2000–2002. Teoksessa R. Hjerppe, A. Kangasharju & R. Vuorento (toim.). Kunnalliset palvelut – Terveysten- ja vanhustenhuollon tuottavuus. VATT, 37. Helsinki.
- Laine, J. (2005a). Henkilöstö ja työvoiman käyttö pitkäaikaishoidossa. Teoksessa A. Noro, H. Finne-Soveri, M. Björkgren & P. Vähäkangas (toim.). Ikääntyneiden laitoshoidon laatu ja tuottavuus - RAI-järjestelmä vertailukehittämisessä. Stakes, M 205. Helsinki.
- Laine, J. (2005b). Pitkäaikaishoidon tuottavuus ja taloudellinen arviointi. Teoksessa A. Noro, H. Finne-Soveri, M. Björkgren & P. Vähäkangas (toim.). Ikääntyneiden laitoshoidon laatu ja tuottavuus - RAI-järjestelmä vertailukehittämisessä. Stakes, M 205. Helsinki.
- Laine, J., Linna, M., Häkkinen, U. & Noro, A. (2005). Measuring the Productive Efficiency and Clinical Quality of Institutional Long-Term Care for the Elderly. *Health Economics* 14(3): 245–256.
- Laine, J., Noro, A. & Finne-Soveri, H. (2004). Mitkä yksilön ominaisuudet ja osaston piirteet selittävät potilaiden saamaa hoitoaikaa pitkäaikaishoidossa? Stakes, Aiheita 3. Helsinki.
- Le Grand, J., Propper, C. & Robinson, R. (1992). *The Economics of Social Problems.* Third edition ed. Palgrave: New York.
- Li, T. & Rosenman, R. (2001). Cost Inefficiency in Washington Hospitals: A Stochastic Frontier Approach Using Panel Data. *Health Care Management Science* 4(2): 73–81.
- Linna, M. (1998). Measuring Hospital Cost Efficiency with Panel Data Models. *Health Economics* 7(5): 415–427.
- Linna, M. (1999). *Measuring Hospital Performance: the Productivity, Efficiency and Costs of Teaching and Research in Finnish Hospitals.* Stakes, Research Report 98. Helsinki.

- Linna, M. & Häkkinen, U. (1999). Determinants of Cost Efficiency of Finnish Hospitals: A Comparison of DEA and SFA. Helsinki University of Technology. Systems Analysis Laboratory Research Reports, A 78. Helsinki.
- Linna, M., Nordblad, A. & Koivu, M. (2003). Technical and cost efficiency of oral health care provision in Finnish health centres. *Social Science & Medicine* 56(2): 343–353.
- Loeb, J. M. (2004). The current state of performance measurement in health care. *International Journal for Quality in Health Care* 16(Supplement I): i5–i9.
- Luoma, K., Järviö, M.-L., Suoniemi, I. & Hjerpe, R. T. (1996). Financial incentives and productive efficiency in Finnish health centres. *Health Economics* 5(5): 435–445.
- Maddala, G. (1992). Introduction to Econometrics. Second edition ed. Prentice-Hall, Inc.
- Mainz, J. (2004). Quality indicators: essential for quality improvement. *International Journal for Quality in Health Care* 16(Supplement I): i1–i2.
- Makkonen, R. & Inkinen, T. (2001). Terveystienhuollon ja sosiaalihuollon laatua käsittelevää kotimaista kirjallisuutta. Stakes, Helsinki.
- McGuire, A. (1985). The theory of the hospital: A review of the models. *Social Science & Medicine* 20(11): 1177–1184.
- McGuire, T. (2000). Physician Agency. In A. J. Culyer & J. P. Newhouse (Eds.), *Handbook of Health Economics* Elsevier Science B V, Volume 1 ed.
- Meeusen, W. & van den Broeck, J. (1977). Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error. *International Economic Review* 18(2): 435–444.
- Mitchell, W., Venkatraman, A., Banaszak-Holl, J., Baum, J. A. C. & Berta, W. (2003). The Commercialization of Nursing Home Care: Does For-Profit Efficiency Mean Lower Quality or do Corporations Provide the Best of Both Worlds. The Fuqua School of Business, Duke University, Durham.
- Newhouse, J. (1970). Toward a Theory of Nonprofit Institutions: An Economic Model of a Hospitals. *American Economic Review* 60(1): 64–74.
- Newhouse, J. (1994). Frontier Estimation: How Useful a Tool for Health Economics? *Journal of Health Economics* 13(3): 317–322.
- Niskanen, W. (1968). Non Market Decision Making - The peculiar economics of bureaucracy. *American Economic Review* 58(2): 293–305.
- Noro, A. (2005). Asiakasrakenne pitkäaikaisessa laitoshoidossa. Teoksessa A. Noro, H. Finne-Soveri, M. Björkgren & P. Vähäkangas (toim.), *Ikääntyneiden laitoshoidon laatu ja tuottavuus - RAI-järjestelmä vertailukehittämisessä* Stakes, M 205. Helsinki.
- Noro, A., Finne-Soveri, H., Björkgren, M., Häkkinen, U., Laine, J., Vähäkangas, P., et al. (2001). RAI-tietojärjestelmän käyttöönotto ja pitkäaikaishoidon benchmarking. Stakes, Aiheita 17. Helsinki.
- Nyman, J. & Bricker, D. L. (1989). Profit Incentives and Technical Efficiency in the Production of Nursing Home care. *The Review of Economics & Statistics* 71(4): 586–594.
- O'Hagan, A. & Luce, B. (2003). Bayesian Statistics in Health Economics and Outcomes Research. MEDTAP International, Inc.
- Okun, A. (1975). Equality and Efficiency – The big trade off. The Brookings Institution: Washington.
- Øvretveit, J. (1998). Evaluating Health Interventions. Open University Press, Buckingham.
- Ozcan, Y. A., Wogen, S. E. & Mau, L. W. (1998). Efficiency evaluation of skilled nursing facilities. *Journal of Medical Systems* 22(4): 211–224.
- Parkkinen, P. (2002). Suomi selviää vanhuusmenoista – mutta entä muuttotappioalueet? *Yhteiskuntapolitiikka*(4): 346–351.
- Parkkinen, P. (2004a). Hoiva- ja hoitopalvelumenot tulevaisuudessa. VATT, 326. Helsinki.
- Parkkinen, P. (2004b). Verotulot lakisääteisistä eläkemenoista sekä julkisista terveydenhuolto- ja sosiaalipalvelumenoista. *Yhteiskuntapolitiikka*(5): 459–471.
- Pekkarinen, L., Sinervo, T., Elovainio, M., Noro, A., Finne-Soveri, H. & Laine, J. (2004). Asiakkaiden toimintakyky, työn organisointi ja henkilöstön hyvinvointi vanhusten pitkäaikaisessa laitoshoidossa. Stakes, Aiheita 11. Helsinki.
- Propper, C. (1996). Market structure and prices: The responses of hospitals in the UK National Health Service to competition. *Journal of Public Economics* 61: 307–335.
- Rantz, M. J., Petroski, G. F., Madsen, R. W., Scott, J., Mehr, D. R., Popejoy, L., et al. (1997). Setting Thresholds for MDS Quality Indicators for Nursing Home Quality Improvement Reports. *Journal on Quality Improvement* 23(11): 602–611.

- Rasbash, J., Browne, W., Goldstein, H., Yang, M., Plewis, I., Healy, M., et al. (2001). A user's guide to MLwiN. Centre for Multilevel Modelling. Institute of Education. University of London: London.
- Reifschneider, D. & Stevenson, R. (1991). Systematic Departures From the Frontier: A Framework for the Analysis of Firm Inefficiency. *International Economic Review* 32(3): 715–723.
- Rice, N. & Jones, A. (1997). Multilevel Models and Health Economics. *Health Economics* 6(6): 561–575.
- Rice, N. & Leyland, A. (1996). Multilevel models: applications to health data. *Journal of Health Services Research & Policy* 1(3): 154–164.
- Roemer, R. & Shain, M. (1959). Hospitals costs relate to the supply of beds. *The Modern Hospital* 92(71–73).
- Rosko, M. (2001). Cost Efficiency of US Hospitals: A Stochastic Frontier Approach. *Health Economics* 10(6): 539–551.
- Rosko, M. (2004). Performance of U.S. Teaching Hospitals: A Panel Analysis of Cost Inefficiency. *Health Care Management Science* 7(1): 7–16.
- Rosko, M., Chilingirian, J., Zinn, J. & Aaronson, W. (1995). The Effects of Ownership, Operating Environment, and Strategic Choices on Nursing Home Efficiency. *Medical Care* 33(10): 1001–1021.
- Sefton, T., Byford, S., McDaid, D., Hills, J. & Knapp, M. (2002). Making the most of it. Economic evaluation in the social welfare field. Joseph Rowntree Foundation: York.
- Sexton, T. R., Leiken, A. M., Sleeper, S. & Coburn, A. F. (1989). The impact of prospective reimbursement on nursing home efficiency. *Medical Care* 27(2): 154–163.
- Shepard, R. (1970). Theory of cost and production functions. Princeton: Princeton University Press.
- Simar, L. & Wilson, P. W. (2003). Estimation and Inference in Two-Stage, Semi-Parametric Models of Production Processes. Unpublished manuscript January 2003.
- Sintonen, H., Pekurinen, M., Jämsen, R., Linnakko, E. & Vinni, K. (1987). Terveystaloustiede. WSOY: Juva.
- StatFin. (2004, 19.1.2004). *StatFin*, 2004, from <http://statfin.stat.fi/statweb>
- Street, A. (2003). How much confidence should we place in efficiency estimates? *Health Economics* 12(11): 895–907.
- SVT. (2003). Care and Services for Older People 2002. Finnish Official Statistics, Social Security: I. Helsinki.
- Syvänen, S. (2003). Työn paineet ja puuttumattomuuden kustannukset. Tutkimus sisäisen tehottomuuden lähteistä ja vaikutuksista, esimerkikikohteena kuntien sosiaalitoimen vanhuspalveluja tuottavat työyhteisöt. Tampereen yliopisto, Acta Universitatis Tamperensis 942. Tampere.
- The Hospital Report Research Collaborative and the Ontario Hospital Association and the Government of Ontario (2001). Hospital report 2001 – Complex continuing care. Toronto, Canada.
- Vaarama, M., Luomahaara, J., Peiponen, A. & Voutilainen, P. (2001). Koko kunta ikääntyneiden asialle. Näkökulmia ikääntyneiden itsenäisen selviytymisen sekä hoidon ja palvelun kehittämiseen. Stakes, Raportteja 259. Helsinki.
- Vaarama, M. & Voutilainen, P. (2002). Kaksi skenaariota vanhusten hoivapalvelujen kehityksestä ja resurssitarpeista ajalla 1999–2030. *Yhteiskuntapolitiikka*(4): 352–363.
- Wang, H.-J. & Schmidt, P. (2002). One-Step and Two-Step Estimation of the Effects of Exogenous Variables on Technical Efficiency Levels. *Journal of Productivity Analysis* 18(2): 129–144.
- Weech-Maldonado, R. J., Neff, G. & Mor, V. (2003). Does quality of care lead to better financial performance? The case of the nursing home industry. *Health Care Management Review* 28(3): 201–216.
- Vitaliano, D. & Toren, M. (1994). Cost and Efficiency in Nursing Homes: A Stochastic Frontier Approach. *Journal of Health Economics* 13(3): 281–300.
- VM. (2003). Tuottavuuden hankesuunnitelma. Julkisen hallinnon ja palvelutuotannon tuottavuuden toimenpideohjelma. Valtionvarainministeriö, Helsinki.
- Worthington, A. C. (2004). Frontier Efficiency Measurement in Health Care: A Review of Empirical Techniques and Selected Applications. *Medical Care Research and Review* 61(2): 135–170.
- Zimmermann, D. R., Karon, S. L. & Arling, G. e. a. (1995). Development and Testing of Nursing Home Quality Indicators. *Health Care Finance Review* 16(4): 107–140.
- Zuckerman, S., Hadley, J. & Iezzoni, L. (1994). Measuring hospital efficiency with frontier cost functions. *Journal of Health Economics* 13(3): 255–280.

Tutkimuksessa käytetyt kliiniset laatuindikaattorit ja niiden laskenta MDS 2.0-lomakkeesta

Aihe	Kuvaus	Muuttujat MDS 2.0-ohjelmalla	Riskiryhmien laskenta
I ONNETTOMUUDET			
1a. Mikä tahansa onnettomuus (e)	Jaettava: Ne asiakkaat, joilla merkitty onnettomuuksia tai kaatumisia. Jakaja: Osaston asiakkaat viimeisimmässä arvioinnissa.	M4_a=1 tai m4_b=1 tai j4_c=1	Ei
1b. Uudet murtumat (i)	Jaettava: Ne asiakkaat, joilla uusi murtuma viimeisimmässä arvioinnissa. Jakaja: Asiakkaat, joilla ei ollut murtumaa viimeisimmässä arvioinnissa	Viimeisin arvio Uusi lonkkamurtuma: j4_c=1 tai uudet murtumat: j4_d=1	Ei
2. Kaatuilu (e)	Jaettava: Asiakkaat, joilla kaatumisia viimeisimmässä arvioinnissa. Jakaja: Asiakkaat viimeisimmässä arvioinnissa.	Kaatonut viimeisen 30 päivän aikana, j4_a=1	Ei
II KÄYTÖS JA TUNNE-ELÄMÄ			
1. Käyttösoire, joka vaikuttaa muihin (e)	Jaettava: Asiakkaat, joilla käyttöoireita, jotka vaikuttavat muihin viimeisimmässä arvioinnissa. Jakaja: Asiakkaat viimeisimmässä arvioinnissa.	Muihin vaikuttavat käyttöoireet Verbaalisesti aggressiivinen: e4ba_ > 0 tai fyysinen pahoinpitely: e4ca_ > 0 tai sosiaalisesti sopimaton/häiritsevä käyttäytyminen: e4da_ > 0	<i>Matala riski:</i> Ne asiakkaat, jotka eivät kuulu korkeaan riskiin viimeisimmässä arvioinnissa. <i>Korkea riski:</i> Jos asiakkaalla kognitiivinen vajaus viimeisimmässä arvioinnissa tai [skitsofrenia (i3a2_-i3e2_ on F20.0–F20.9 tai i1_gg=1) tai maanisdepressiivisyys (i3a2_-i3e2_ on F30–F39 tai i1_ff=1)]

Aihe	Kuvaus	Muuttujat MDS 2.0-ohjelmalla	Riskiryhmien laskenta
2. Masennusdiagnoosi tai -oireet	Jaettava: Asiakkaat, joilla on masennusoireita tai masennusdiagnoosi viimeisimmässä arvioinnissa. Jakaja: Asiakkaat viimeisimmässä arvioinnissa.	Masennusoireet 1. oire = ahdinko: e1a_=1,2 2. oire = ahdistus/vetäytyminen: toistuvat fyysiset eleet e1_n=1,2, tai vastustaa hoitoa e4ea_=1,2,3, tai vetäytyy aktiviteeteistaan e1_o=1,2, tai vähentynyt sosiaalinen kanssakäyminen: e1_p=1,2 3. oire = huonotuulusuus aamuisin: e1j_=1,2, tai ei hereillä koko päivää: ni_d=1, tai hereillä yhden jakson päivässä tai vähemmän, mutta ei tajuton: n1_a, n1_b, n1_cf1 ja b1__=0 4. oire = toistuvasti arvelee kuolevansa: e1g_=1,2 5. oire = painon lasku: k3a_=1	Ei
3. Masennuksen oireet – ei hoitoa (e)	Jaettava: Asiakkaat, joilla on masennusoireita viimeisimmässä arvioinnissa ja ei ole hoitoa. Jakaja: Asiakkaat viimeisimmässä arvioinnissa.	Masennus, ks. edellinen ja ei antidepressanttia: o4c_=0 ja ei psykologista terapiaa: p1bea_=0	Ei

III KLIININEN HOITO

1. Käytössä 9 erilaista tai useampi lääke (e)	Jaettava: Asiakkaat, joilla oli 9 erilaista tai useampi lääke. Jakaja: Kaikki asiakkaat viimeisimmässä arvioinnissa.	<i>Viimeisin arviointi:</i> O1 (Lääkkeiden lukumäärä = 9 tai enemmän)	Ei
---	---	---	----

IV KOGNITIO

1. Aivojen vajaatoiminta (i)	Jaettava: Asiakkaat, joilla äskettäin aivojen vajaatoiminta viimeisimmässä arvioinnissa. Jakaja: Asiakkaat, joilla ei aivojen vajaatoimintaa edellisessä arvioinnissa.	<i>Viimeisin arviointi:</i> Aivojen vajaatoiminta. <i>Aiempi arviointi:</i> Ei ole aivojen vajaatoimintaa.	Ei
------------------------------	---	---	----

Aihe	Kuvaus	Muuttujat MDS 2.0-ohjelmalla	Riskiryhmien laskenta
V PIDÄTYSKYKY			
1. Virtsan/ulosteen pidätyskyvyttömyys (e)	<p>Jaettava: Asiakkaat, jotka usein tai täysin pidätyskyvyttömiä viimeisimmässä arvioinnissa.</p> <p>Jakaja: Kaikki muut asiakkaat paitsi poissuljetut.</p>	<p><i>Viimeisin arviointi:</i> Virtsan pidätyskyky: H1b_ = 3,4 TAI Ulosteen pidätyskyky: H1a_ = 3,4</p> <p>POISSULJETTU: Tajuttomat asiakkaat: b1__=1 TAI asiakkaalla katetri: h3_d TAI asiakkaalla avanne: h3_i viimeisimmässä arvioinnissa</p>	<p><i>Korkea riski:</i> Vaikea kognitiivinen vaje TAI päivittäisten toimintojen suori-tuskyvyssä täysin autettava (g1aa_, g1ba_, g1ea_=4) viimeisimmässä arvioinnissa</p> <p><i>Matala riski:</i> Kaikki muut viimeisimmässä arvioinnissa</p>
2. Virtsan/ulosteen ajoit-tainen pidätyskyvyttömyys (e) – ei suolen/rakon koulutusohjel-maa	<p>Jaettava: Asiakkaat, joilla ei ole koulutusohjelmaa.</p> <p>Jakaja: Asiakkaat, jotka toisi-naan tai usein pidätyskyvyttömiä (rakko tai suoli) viimeisimmässä arvioinnissa.</p>	<p><i>Viimeisin arviointi:</i> Ei ohjel-moitua wc-opetusta ja rakon koulutusohjelmaa: Toisinaan tai usein pidätyskyvytön rakon tai suolen toiminnassa: h1a_ =2,3 TAI h1b_ =2,3</p>	Ei
3. Katetri (e)	<p>Jaettava: Katetri viimeisim-mässä arvioinnissa.</p> <p>Jakaja: Kaikki asiakkaat vii-meisimmässä arvioinnissa.</p>	<p><i>Viimeisin arviointi:</i> Katetri: h3_d=1</p>	<p><i>Korkea riski:</i> Tajuton: b1__=1, tai hemiplegia: i1_v=1, tai tet-raplegia: i1_z=1, tai multippeli skleroosi: i1_w=1, tai täysin autet-tava liikkumi- sessa TDMAL=1: g1aa_=4 ja g1ba_=4 ja g1ea=1</p> <p><i>Matala riski:</i> Ei ole täysin autettava liikkumi- misessä: bi__=0 ja i1_v=0 ja i1_z=0 ja i1_w=0 ja TDMAL=0</p>
4. Ulosteen pakkautumi-nen (e)	<p>Jaettava: Asiakkaat, joil-la ulosteen pakkautumista vii-meisimmässä arvioinnissa.</p> <p>Jakaja: Kaikki asiakkaat vii-meisimmässä arvioinnissa.</p>	<p><i>Viimeisin arviointi:</i> Ulosteen pakkautuminen: h2_d=1</p>	Ei

Aihe	Kuvaus	Muuttujat MDS 2.0-ohjelmalla	Riskiryhmien laskenta
VI INFECTIOIDEN HALLINTA			
1. Virtsainfektiot (e)	Jaettava: Asiakkaat, joilla oli VTI. Jakaja: Kaikki asiakkaat viimeisimmässä arvioinnissa.	<i>Viimeisin arviointi:</i> I2j (VTI) rastitettu.	Ei
VII RAVINTO JA SYÖMINEN			
1. Painon lasku (e)	Jaettava: Ne asiakkaat, joilla viimeisimmässä arvioinnissa oli 5 %:n painon lasku viimeksi kuluneen 30 vrk:n aikana tai 10 % tai enemmän viimeksi kuluneen 10 päivän aikana. Jakaja: Kaikki asiakkaat viimeisimmässä arvioinnissa.	<i>Viimeisin arviointi:</i> K3a=1 (painon lasku).	Ei
2. Letkuravitus (e)	Jaettava: Letkuravitus viimeisessä arvioinnissa. Jakaja: Kaikki asiakkaat viimeisimmässä arvioinnissa.	<i>Viimeisin arviointi:</i> K5b (letku).	Ei
3. Kuivuma (e)	Jaettava: Asiakkaat, joilla on kuivuma. Jakaja: Kaikki asiakkaat viimeisimmässä arvioinnissa.	<i>Viimeisin arviointi:</i> J1c rästetty tai kuivuman ICD-10-diagnoosi annettu.	Ei
VIII TOIMINTAKYKY			
1. Vuoteeseen hoidettavat potilaat (e)	Jaettava: Vuoteeseen hoidettavat potilaat viimeisimmässä arvioinnissa. Jakaja: Kaikki asiakkaat viimeisimmässä arvioinnissa.	<i>Viimeisin arviointi:</i> Pääasiallisesti vuodepotilas: g6_a=1	Ei

Aihe	Kuvaus	Muuttujat MDS 2.0-ohjelmalla	Riskiryhmien laskenta
2. Toimintakyvyn lasku koskien pitkään säilyviä taitoja (i)	<p>Jaettava: Asiakkaat, joilla päivittäisten toimintojen suorituskyvyn laskua viimeisen ja sitä edellisen mittauksen välillä.</p> <p>Yhden tason lasku kahdessa tai useammassa TAI kahden tason lasku yhdessä tai useammassa.</p> <p>Jakaja: Kaikki viimeiseen ja sitä edelliseen mittaukseen osallistuneet asiakkaat. (Päiti ne, joiden toimintakyky ei voi enää laskea, koska olivat edellisessä mittauksessa täysin autettavia tai olivat tajuttomia.)</p>	<p>Ainakin YHDEN tason lasku kahdessa tai useammassa seuraavista: liikkuminen sängyssä (g1aa_), siirtyminen (g1ba_), ruokailu (g1ha_), wc:n käyttö (g1ia_)</p> <p><i>Edellinen Viimeisin mittaus mitta</i></p> <p>1,2,3 tai 4 2,3 tai 4 3 tai 4 4 TAI</p> <p>Ainakin KAHDEN tason lasku yhdessä tai useammassa seuraavista: liikkuminen sängyssä (g1aa_), siirtyminen (g1ba_), ruokailu (g1ha_), wc:n käyttö (g1ia_)</p> <p><i>Edellinen Viimeisin mittaus mitta</i></p> <p>2,3,4 3,4 4</p> <p>Huom. Arvo 8 on = puuttuva määriteltäessä muutosta ADL:ssä.</p> <p>POISSULJETTU: Asukkaat, joiden toimintakyky ei voi enää laskea. (g1aa_-aj – kaikki = 4 tai 8) TAI tajuttomia (b1_=1) edellisessä mittauksessa.</p>	On
3. Nivelten toiminnallisen liikelaa-juuden mene-tyt (i)	<p>Jaettava: Asiakkaat, joilla nivelten liikelaa-juuden rajoitukset ovat lisääntyneet viimeisen ja sitä edellisen mittauksen välillä.</p> <p>Jakaja: Kaikki asiakkaat viimeisessä ja sitä edellisessä mittauksessa huomioiden poissuljetut.</p>	<p>Nivelten liikelaa-juuden rajoitukset (g4aa_ - af >0) viimeisimmässä arvioinnissa ovat suuremmat kuin nivelten liikelaa-juuden rajoitukset sitä edellisessä mittauksessa.</p> <p><i>Viimeisin Edellinen mittaus mitta</i></p> <p>Summa g4aa_-af > Summa g4aa_-af</p> <p>POISSULJETTU: asiakkaat, joilla suurimmat mahdolliset nivelten liikelaa-juuden rajoitukset. (Summa g4aa_ - af = 12 edellisessä mittauksessa.)</p>	Ei

Aihe	Kuvaus	Muuttujat MDS 2.0-ohjelmalla	Riskiryhmien laskenta
4. Harjoittelun/taitojen parantamisen/ liikelaajuuden ylläpitämisen puute liikuntarajoitteisilla asukkailla (e)	Jaettava: Ei harjoittelua eikä taitojen parantamista. Jakaja: Liikuntarajoitteiset.	<i>Liikuntarajoitteiset:</i> G1aa_ = 2,3,4 TAI g1ba_ = 2,3,4 TAI g1ea_ = 2,3,4 JA jos ei hoitajien antamaa kuntoutusta ja ylläpitohoitoa: p3d_=0 ei vuoteessa liikkumisen opastusta ja harjoittelua JA p3e_=0 ei siirtymisen opastusta ja harjoittelua JA p3f_=0 ei kävelemisen opastusta ja harjoittelua JA p3a_=0 ei passiivista liikehoitoa JA p3b_=0 ei aktiivista liikehoitoa.	Ei
IX PSYYKENLÄÄKKEIDEN KÄYTTÖ			
1. Antipsykoottisten lääkkeiden käyttö ilman psykoottisia (tai vast.) oireita (e)	Jaettava: Antipsykoottisia lääkkeitä viimeisimmässä arvioinnissa käyttävät asiakkaat. Jakaja: Kaikki muut asiakkaat viimeisimmässä arvioinnissa paitsi psykoottiset tai vastaavat.	<i>Viimeisin arviointi:</i> Antipsykoottiset (o4a_ ³ 1) POISSULJETTU: Asiakkaat, joilla yksi tai useampia psykoottisia oireita (i3=F20-F20.9; F22-F22.9, F23.0, F24, F25.00-25.9, F28, F29 tai skitsofrenia todettu i1_gg) TAI Touretten oireyhtymä (i3=F95.2) TAI Huntingtonin tauti (i3=G10) viimeisimmässä arvioinnissa tai viimeisimmässä täydessä arvioinnissa; TAI todettu harhaisuutta (j1_i) viimeisimmässä arvioinnissa.	<i>Korkea riski:</i> Kognitiivinen heikennys JA käytösongelmat viimeisimmässä mittauksessa. <i>Matala riski:</i> Kaikki muut viimeisimmässä mittauksessa.

Aihe	Kuvaus	Muuttujat MDS 2.0-ohjelmalla	Riskiryhmien laskenta
2. Rauhoittavien ja unilääkkeiden käyttö (e)	<p>Jaettava: Asiakkaat, jotka saaneet anksiolyyttiä tai unilääkettä viimeisimmässä arvioinnissa.</p> <p>Jakaja: Kaikki muut asiakkaat viimeisessä mittauksessa paitsi psykoottiset tai vastaavat.</p>	<p><i>Viimeisin arviointi:</i> Anksiolyytti/unilääke (o4b_ tai o4d_31)</p> <p>POISSULJETTU:</p> <p>Asiakkaat, joilla yksi tai useampia psykoottisia oireita (i3=F20-F20.9; F22-F22.9, F23.0, F24, F25.00, F28, F29 tai skitsofrenia todettu i1_gg)</p> <p>TAI Touretten oireyhtymä (i3=F95.2)</p> <p>TAI Huntingtonin tauti (i3=G10) viimeisimmässä arvioinnissa tai viimeisimmässä täydessä arvioinnissa;</p> <p>TAI todettu harhaisuutta (j1_i) viimeisimmässä arvioinnissa.</p>	Ei
3. Unilääkkeiden käyttö useamman kuin kahdesti viikossa (e)	<p>Jaettava: Asiakkaat, jotka saaneet unilääkettä useammin kuin kahdesti viikossa viimeisimmässä arviossa.</p> <p>Jakaja: Kaikki asiakkaat viimeisessä mittauksessa.</p>	<p><i>Viimeisin arviointi:</i> Unilääkkeiden käyttö useammin kuin kahdesti viikossa (o4d_>2)</p>	Ei
X ELÄMÄN LAATU			
1. Päivittäinen liikkumista rajoittavien välineiden käyttö (e)	<p>Jaettava: Asiakkaat, joiden liikkumista oli rajoitettu päivittäin viimeisessä arviossa.</p> <p>Jakaja: Kaikki asiakkaat viimeisimmässä arvioinnissa.</p>	<p><i>Viimeisin arviointi:</i> Päivittäisen "sitominen" = liikkumisen rajoittaminen P4c, d tai e=2.</p>	Ei
2. Vähän tai ei lainkaan aktiiviteettia (e)	<p>Jaettava: Asiakkaat, joilla oli vähän tai ei lainkaan aktiiviteettejä viimeisessä arviossa.</p> <p>Jakaja: Kaikki asiakkaat viimeisimmässä arvioinnissa (poissuljettu koomassa olevat).</p>	<p><i>Viimeisin arviointi:</i> N2=2 tai 3, vähän tai ei lainkaan aktiiviteettia.</p> <p>POISSULJETTU: Tajuttomat asiakkaat B1=1.</p>	Ei

Aihe	Kuvaus	Muuttujat MDS 2.0-ohjelmalla	Riskiryhmien laskenta
XI IHON HOITO			
1. 1–4 asteen painehaavat (e)	Jaettava: Asiakkaat, joilla painehaavoja (aste 1–4) viimeisimmässä arvioinnissa. Jakaja: Kaikki asiakkaat viimeisimmässä arvioinnissa.	<i>Viimeisin arviointi:</i> Painehaava: m2a_=1,2,3,4 tai i3=L89.	<i>Korkea riski:</i> Liikkuminen sängyssä (g1aa_=3,4) tai siirtyminen (g1ba_=3,4), TAI kooma (b1__=1), TAI aliravitsemus (i3=E40, tai E41, tai E42, tai E43, tai E44) TAI sairauden terminaalivaihe (j5_c) TAI hemiplegia (i1_v=1) TAI tetraplegia (i1_z=1) TAI turvotus (j1_g=1) TAI iho tunnoton paineelle tai kivulle (m4_e=1) viimeisimmässä arvioinnissa. <i>Matala riski:</i> Kaikki muut viimeisimmässä arvioinnissa.

Lähde: Noro et al. 2001.

The association between quality of care and technical efficiency in long-term care

JUHA LAINE¹, U. HARRIET FINNE-SOVERI¹, MAGNUS BJÖRKGREN², MIIKA LINNA¹, ANJA NORO¹
AND UNTO HÄKKINEN¹

¹Centre for Health Economics at STAKES – CHESS, Helsinki, Finland, ²Chydenius Institute, University of Jyväskylä, Kokkola, Finland

Abstract

Objective. To analyse the association between quality of care and technical (productive) efficiency in institutional long-term care wards for the elderly.

Setting. One hundred and fourteen public health centre hospitals and residential homes in Finland.

Study design. Wards were divided into two categories according to their rank in the quality distribution, considering 41 quality variables separately. The technical efficiency scores of the good- and poor-quality groups were compared using cross-sectional data.

Methods. Data envelopment analysis was used for calculating technical efficiency. The Mann–Whitney test and correlation coefficients were used to explore the association between quality and efficiency.

Results. The wards where quality indicators indicated less pro-active (passive) nursing practice and more dependent patients—for instance, in terms of very high prevalence of bedfast residents or very high prevalence of daily physical restraints—performed more efficiently than the comparison group.

Conclusion. The results suggest that an association may exist between technical efficiency and unwanted dimensions of quality. Hence, the efficiency and quality of care are essential aspects of management and performance measurement in elderly care.

Keywords: data envelopment analysis, long-term care, quality, technical efficiency

The association between economic performance and quality of care is an essential aspect of the production of health services. According to Newhouse [1], quantity and quality are two commodities to which the decision-maker can allocate resources, and that a quantity–quality trade-off is present when resources are constrained. It is appealing to think that increasing quality may require additional labour and capital resources, whilst a tendency towards efficiency improvements and cost containment can lead to a poorer performance in quality. On the other hand, better quality can be associated with better economic performance and lower production costs, i.e. better efficiency [2]. Deterioration in quality can cause voiced complaints (voice) and possible withdrawal (exit) of customers, which force the management of the unit to react and to prevent economic loss [3]. If there is an association between economic performance and quality of care in health services, it definitely has implications relevant to policy and management.

Technical (productive) efficiency is one possible way to evaluate the economic performance of the units. It is a measure of the proportion of input resources (e.g. labour) to the

output (e.g. in-patient days). When units are compared, efficient units produce the same output with fewer inputs than inefficient units, or alternatively, produce more output with the same input as inefficient units [4]. Björkgren *et al.* [5] and Laine *et al.* [6] have shown that the average technical inefficiency is high in Finnish long-term care units. The variation in clinical quality of care and staff per resident ratio within these units has also been shown to be marked [6,7]. Therefore, particular attention should be paid to these differences in performance and the potential association between efficiency and quality of care. Differences in efficiency may result in differences in the quality of care produced and the number of nursing staff.

Defining and measuring the quality of care in long-term care facilities is a multidimensional and complex issue with several pitfalls [8]. Firstly, the majority of the residents suffer from cognitive decline [7] and thus, ‘the voice of the client’ might not reach the management. Decline in cognition may also indicate that residents cannot necessarily evaluate quality, which signifies an asymmetry in information between residents and management. Secondly, Kane [9] argues that the

Address reprint requests to Juha Laine, STAKES, Lintulahdenkuja 4, PO Box 220 FIN-00531 Helsinki, Finland. E-mail: juha.laine@stakes.fi

goals of care among the residents are far from uniform. A minority of this population consists of individuals with good prognosis and discharge potential. Kane [9] further states that high-quality care, in these settings, may also mean that the patient does less poorly than otherwise expected.

There are few studies of the relationship between staffing and quality in elderly care that show consistently that better staffing is associated with better quality of care [10]. In Finland the present discussion concerning public services is related mostly to productive efficiency or productivity. Unfortunately, there are not yet clear messages or tools for policy-makers, since the empirical evidence on the association between efficiency and quality in elderly care seems to be scarce and somewhat contradictory [11–13].

The aim of this study was to consider the association between efficiency and quality of care in long-term care wards in Finland. Due to the multidimensional nature of quality, 41 quality indicators were used. In the first stage, instead of using traditional staffing figures—such as the staff per resident ratio—we calculated the productive efficiency of the units. A unit's efficiency score comprises the following essential components: the number of staff, the amount of capital, the total number of in-patient days, and the case-mix. In the second stage, units were divided into poor- and good-quality groups and the efficiency scores of these groups were compared.

Materials and methods

Data

The observation unit in this study was a long-term care ward. For the study we needed information on the staff, the number of the beds, the case-mix, and the quality of care in the wards. Complete research data pertaining to 114 wards in 41 facilities were gathered from 1 January 2002 to 31 December 2002 from several sources. There were approximately 3300 patients on the wards.

Client-level assessments have been done using the Minimum Data Set 2.0 (MDS 2.0) questionnaire for nursing facilities [14]. In Finland, each patient is assessed by the staff with the assistance of the resident's relatives both at admission and semi-annually, or when a distinct change in the patient's condition takes place. Resource Utilization Groups (RUG-III) were used to represent the case-mix [15]. It can be constructed from the MDS together with clinical quality indicators indicating prevalence and incidence of adverse outcomes and care processes [14,16,17].

The RUG-III classification is internationally the most widely used and validated method for case-mix adjustment [18–20]. In the USA, the RUG-III system forms the basis of prospective payment of publicly reimbursed nursing home care. RUG-III is based on information in the Resident Assessment Instrument (RAI) and its MDS, such as patient's physical functioning and cognition. The RAI is a standardized assessment instrument for improving care planning and quality of care. In Finland, a 22-group version of RUG-III classification is used [19]. In the RUG-III/22 patients are placed

into one of the seven clinical categories, which are divided into 22 subcategories. The subcategories measure how much labour resources a patient needs per day. Each of the subcategories is cost weighted, and the cost weight denotes the average proportional staff time per 24-hour period per patient. The cost weight for the average patient is set at 1.0. The cost weights for different patient groups vary from 0.48 to 2.52 in RUG-III/22, which denotes that there is a six-fold difference in average nursing time per 24-hours between patients. In this study we used a case-mix index, which was aggregated for the ward level from individual-level data.

Clinical quality indicators at ward level were calculated as mean values of measurements made during two MDS assessment periods (cross-sections) in 2002, in accordance with Zimmermann *et al.* [17]. The pooling together (calculated averages) of two cross-sections ensures more consistent quality measures and diminishes measurement errors and randomness. The prevalence and incidence of the quality indicators in relative terms were calculated by including all patients on the ward, or including only low-risk patients, or only high-risk patients. These clinical quality indicators are available for routine reporting from our RAI benchmarking database. A questionnaire filled in by the head nurses provided data on staff resources and structural quality measures, such as the number of staff per resident and the proportion of single rooms. The descriptive statistics of the quality variables are given in Table 1.

An average number of three different full-time equivalent staff categories and capital (unit size) were used as input. The in-patient days of the wards were obtained from the finance offices. We used the RUG-III/22 classification for case-mix adjustment [19], which was applied to the output variable, i.e. to the number of in-patient days. By applying the case-mix adjustment the wards are comparable in terms of the output measure, which is crucial in efficiency measurement. The descriptive statistics for the output and input variables are given in Table 2.

Statistical methods

Calculating efficiency. The non-parametric approach, data envelopment analysis, is a traditional method for calculating (in)efficiency [4,21]. Technical inefficiency in this context means that a technically efficient producer could produce the same output with less input when compared with the inefficient unit. Alternatively, a technically efficient producer could produce more output with the same input. These correspond to an input- or output-oriented data envelopment analysis model perspective, respectively. An illustrative example of the data envelopment analysis is given in Figure 1.

For example, consider a production technology of two inputs and one output. The units A, B, C, and D in Figure 1 produce the same output with different combinations of the inputs. The efficiency scores are determined by the ratio of the sum of output to the sum of corresponding input. The efficient units in the frontier (A, B, D) have a score of 1.0. The unit C is inefficient, because it uses more input for the same output compared with the other units. The efficiency

Table 1 Descriptive statistics of the clinical and structural quality indicators

Variable	Mean	Standard deviation	Minimum– maximum	85th percentile
Clinical quality indicators indicating adverse care processes and outcomes ^{1,2}				
Prevalence of any injury, %	27	15	0–76	43
Incidence of new fractures, %	1	2	0–16	4
Prevalence of falls, %	11	8	0–48	20
Prevalence of behavioural symptoms affecting others, %	34	15	6–72	51
Prevalence of behavioural symptoms affecting others, low risk, %	21	21	0–100	44
Prevalence of behavioural symptoms affecting others, high risk, %	38	15	0–78	55
Prevalence of diagnosis or symptoms of depression, %	34	16	6–86	49
Prevalence of depression with no treatment, %	14	7	0–54	20
Use of nine or more different medications, %	35	17	0–82	54
Incidence of cognitive impairment, %	14	15	0–100	25
Prevalence of bladder or bowel incontinence, %	64	17	25–100	83
Prevalence of bladder or bowel incontinence, low risk, %	45	20	13–98	64
Prevalence of bladder or bowel incontinence, high risk, %	83	16	0–94	86
Prevalence of occasional or frequent bladder or bowel incontinence without a toileting plan, %	48	31	0–100	100
Prevalence of indwelling catheters, %	4	5	0–50	9
Prevalence of indwelling catheters, low risk, %	3	6	0–28	9
Prevalence of indwelling catheters, high risk, %	5	8	0–33	14
Prevalence of faecal impaction, %	11	11	0–65	22
Prevalence of urinary tract infections, %	14	8	0–42	23
Prevalence of weight loss, %	7	5	0–28	12
Prevalence of bedfast residents, %	22	20	0–85	46
Incidence of decline in late loss activities of daily living, %	25	15	0–100	36
Incidence of decline in late loss activities of daily living, low risk, %	22	16	0–100	33
Incidence of decline in late loss activities of daily living, high risk, %	41	31	0–100	80
Incidence of decline in range of motion, %	18	11	0–100	25
Incidence of decline in range of motion, low risk, %	16	12	0–67	28
Incidence of decline in range of motion, high risk, %	18	18	0–69	36
Lack of training/skill practice or range of motion for mobility dependent residents, %	27	25	0–95	60
Prevalence of antipsychotic use, in the absence of psychotic and related conditions, %	41	16	0–100	57
Prevalence of antipsychotic use, in the absence of psychotic and related conditions, low risk, %	36	17	7–92	52
Prevalence of antipsychotic use, in the absence of psychotic and related conditions, high risk, %	58	24	0–93	83
Prevalence of anti-anxiety/hypnotic use, %	58	15	12–92	74
Prevalence of hypnotic use more than twice in last week, %	36	14	0–79	49
Prevalence of daily physical restraints, %	19	13	0–59	35
Prevalence of little or no activity, %	63	19	4–100	84
Prevalence of stage 1–4 pressure ulcers, low risk, %	3	6	0–40	8
Prevalence of stage 1–4 pressure ulcers, high risk, %	12	9	0–50	21
Incidence of new pressure ulcers, %	5	5	0–22	11
Structural quality measures ³				
Proportion of registered nurses, %	28	14	6–96	13
Proportion of rooms with own toilet, %	63	42	0–100	5
Proportion of single rooms, %	46	30	0–100	13

¹Low- or high-risk adjusted indicator means that the denominator of the indicator comprises only patients of low or high risk, instead of all patients on the ward.

²Higher prevalence or incidence indicates poorer quality.

³Higher prevalence indicates better quality. The 15th percentile was used instead of 85th percentile for poor-quality threshold.

Table 2 Descriptive statistics of output and input used in the data envelopment analysis

Variable	Description	Mean (standard deviation)	Median (minimum–maximum)
Output			
Case-mix weighted in-patient days	Ward’s in-patient days in 2002 * ward’s average RUG-III/22 case-mix index	11539 (5387)	10810 (2467–30486)
Inputs			
Registered nurses (RNs)	Average number of RNs in wards	5.2 (3.3)	5.0 (0–15)
Licensed practical nurses (LPNs)	Average number of LPNs in wards	9.8 (3.9)	9.0 (0–25)
Aides	Average number of aides in wards	3.2 (2.3)	3.0 (0–12)
Unit size	The number of beds	30.0 (12.0)	28 (8–74)

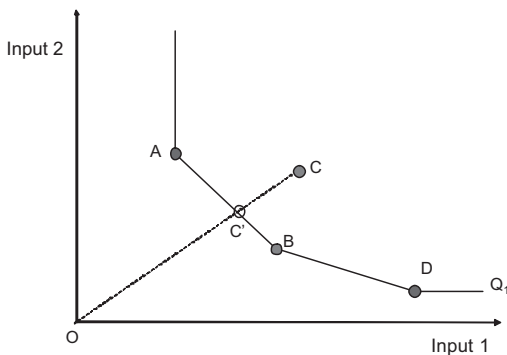


Figure 1 The production of the same quantity of output with different combinations of inputs. Q_1 denotes the best practice production frontier; A, B, and D are efficient units and C' is a reference unit for inefficient unit C; dashed line is a radial measure of inefficiency.

score for unit C is about 0.75 (the ratio OC'/OC), which denotes 25% inefficiency. Unit C could produce its output with 25% less input resources.

The frontier, isoquant Q_1 , in Figure 1 denotes the best practice production frontier. Linear programming is used in the data envelopment analysis to construct an empirical piecewise linear efficient frontier which serves as a reference in the evaluation of efficiency. Data envelopment analysis can also be used for calculating allocative and cost efficiency, which also takes input mix and prices into account. Technical efficiency describes the efficiency of production, which is a necessary but not sufficient precondition for cost efficiency and lowest possible unit costs. In the framework of this study a ward is judged to be efficient if it is operating on the best practice production frontier (see Figure 1). Data envelopment analysis efficiency scores were calculated using the DEAP program written by Tim Coelli. Input price variables (wages) were not available for this study, so an analysis was restricted to exploring technical efficiency, using three categories of labour and capital variables as input, and case-mix adjusted

in-patient days as output. Compared with traditional staffing figures, the efficiency score also takes case-mix and the amount of the output into account.

Quality measures as correlates of inefficiency. We applied a totally non-parametric approach in the analysis of quality measures as correlates of technical inefficiency. We divided wards into two categories according to their rank in the quality distribution, considering all quality variables separately. If a ward was ranked in the zero to 85th percentiles in the quality distribution, it was placed in the group of good or acceptable quality. In contrast, wards that were above the 85th percentile threshold, in which the prevalence and incidence of the specific quality indicator was the highest, were placed in the group of poor quality. There may be severe quality problems in the group above the 85th percentile threshold, and therefore we pay particular attention to these wards in this study.

The traditional Mann–Whitney U-test was used to test whether the technical efficiency scores of the two sampled populations were equivalent in rank. The analysis was conducted with pooled data of health centre hospitals and residential homes ($n = 114$). In the sensitivity analysis we changed the thresholds from 0–85th percentiles (good quality) and above the 85th percentile (poor quality) to 0–75th percentiles (good quality) and above the 75th percentile (poor quality). In addition, we also combined the two groups and tested the correlation between efficiency and continuous quality indicators in the whole sample.

Results

In Figure 2 the Salter diagram illustrates the efficiency distribution. The width of the bars denotes the number of beds in the wards, which varied from 8 to 74. The mean technical efficiency in pooled data was 0.72. Hence, on average, wards could produce in-patient days with 28% less input resources. Four units out of 114 were fully efficient, i.e. efficiency scores were 1.0, which determined the production frontier. There was no clear association between unit size and technical efficiency.

Table 3 presents the results of the analyses tests. The third column denotes the mean rank in efficiency in the comparison

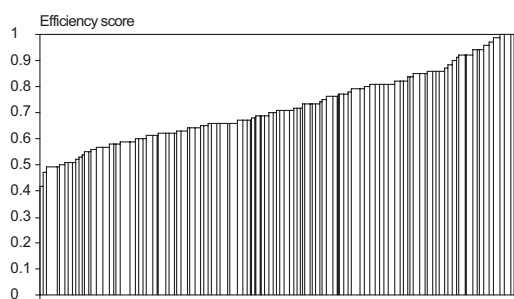


Figure 2 Wards arranged by increasing technical efficiency scores. The y-axis denotes efficiency score, the x-axis denotes wards, and the width of each bar is the ward's size.

groups. In the fourth column the asymptotic significance of the Mann–Whitney test is reported. The correlation coefficients between efficiency score and continuous quality indicators in the pooled data ($n = 114$) are presented in the fifth column. In the last column a positive sign denotes that the wards where the quality seems to be poorest operate more efficiently than wards where quality is good or adequate, according to clinical quality indicators and structural quality measures. A negative sign denotes that poor-quality wards are also less efficient than good-quality wards. The sign $+/-$ denotes no clear association between quality and efficiency.

Positively significant clinical quality indicators in Table 3 are associated with clinically complex and very dependent patients. For instance, the mean rank for bedfast residents was 51 in the good-quality group compared with 92 in the poor-quality group, which indicates clearly that better efficiency is associated with nursing practices which lead to poorer quality in terms of a high number of bedfast residents. However, indicators describing less dependent patients, i.e. patients who walk, who have behavioural problems, or who are more medicated are negatively or non-significantly associated with efficiency. Furthermore, the lower the proportion of registered nurses and proportion of single rooms, the better the efficiency; whereas wards operated more efficiently when they had a higher number of rooms with en-suite toilets. The correlation coefficients between efficiency and quality indicators in the whole sample indicated consistent results with the Mann–Whitney tests. Hence, there also seems to be a linear association between quality indicators and efficiency scores.

In the sensitivity analysis we changed the threshold for poor quality from the 85th percentile to the 75th percentile. The differences in mean ranks between the groups were consistent (the results are not reported in the tables). The only exceptions were a prevalence of depression with no treatment and a proportion of rooms with their own toilet, in which there were no differences in efficiency between the groups. Neither of the correlations were significant.

Discussion and conclusions

There was great variation in efficiency and quality between the wards. Mean technical inefficiency was 28%, which was higher than was estimated in earlier studies [22]. The variation in quality also showed great differences between wards. Our findings showed that in 13 of 41 areas, poor quality was associated with high productive or technical efficiency (see Table 3), indicating also an association between costs, staffing, and quality of care; whereas in six areas the poor quality was related to low efficiency. In the remaining 22 areas no clear association was found.

Efficient wards had quality problems in the areas that indicate time-consuming nursing procedures. These procedures were not in place or they were inadequately performed because of a low staffing ratio. A high prevalence of bedfast residents, the use of physical restraints, a high percentage of inactive patients, and a lack of nursing rehabilitation are all different aspects of mobility restriction, a known risk factor in the disablement process [23]. Use of physical restraints has previously been shown to be associated with low staffing levels in the USA [24]. Grade 1–4 pressure ulcers can be seen as a consequence of both mobility restrictions [25] and incontinence. In addition, a high occurrence of pressure ulcers reflects an absence of or inadequate repositioning programmes in dependent patients. During the disablement process, acceleration of incontinence occurs. A high prevalence of occasional or frequent bladder or bowel incontinence without a toileting plan was associated with high efficiency, though the statistical significance level varied in the non-parametric analysis.

Efficient wards showed a low prevalence of falls. Physical restraints are usually used to prevent falls or injuries [26,27]. Conversely, it has been pointed out that the use of these devices has itself been a source of injuries [26]. In our study, a low prevalence of falls might simply reflect the highest possible level of disability of patients in the efficient wards: bedfast residents are less likely to fall [7].

Low efficiency was found to be associated with poor quality in the areas of disruptive behaviour and the use of antipsychotic medication in the absence of psychotic conditions. The presence of neurodegenerative and/or psychiatric disease is most often the source of disruptive behaviour in these settings [28,29]. In spite of physical disability, wards specializing in dementia care or in psychiatric care—especially those with severe behavioural disturbances as eligibility criteria—are traditionally provided with a higher staff ratio than ordinary wards [30]. In these areas the basic question of quality is whether the use of antipsychotic medication actually improves the individual's quality of life in the long run; a question still open to debate.

Quality measures related to nutrition were not associated with efficiency at all. This might reflect the allocation of nursing time as much as to prioritizing the most important areas: whatever the staffing ratio, issues of feeding and hydration are considered more important than rehabilitation. Neither was the high prevalence of anti-anxiety and hypnotic medication associated with efficiency. One explanation might be the wide use of such medication in all conditions and their unclear relationship with time-consuming care processes.

Table 3 Tests and correlations between efficiency scores and quality variables

Variable	Quality groups	Mean rank in efficiency	Asymptotic sig. ¹	Correlation between efficiency scores and continuous quality indicators in pooled data (<i>n</i> = 114), <i>P</i> values in parentheses	Association between poor quality and high efficiency ²
Prevalence of falls, %	Good quality	60	0.03	−0.37 (<0.001)	−
	Poor quality	41			
Prevalence of behavioural symptoms affecting others, %	Good quality	60	0.09	−0.22 (0.02)	−
	Poor quality	45			
Prevalence of behavioural symptoms affecting others, low risk, %	Good quality	60	0.22	−0.19 (0.04)	−
	Poor quality	48			
Prevalence of behavioural symptoms affecting others, high risk, %	Good quality	60	0.03	−0.26 (0.01)	−
	Poor quality	41			
Prevalence of antipsychotic use, in the absence of psychotic and related conditions, high risk, %	Good quality	62	<0.001	−0.37 (<0.001)	−
	Poor quality	32			
Prevalence of antipsychotic use, in the absence of psychotic and related conditions, low risk, %	Good quality	63	<0.001	−0.27 (<0.001)	−
	Poor quality	29			
Proportion of rooms with own toilet, %	Good quality	60	0.07	−0.01 (0.90)	−
	Poor quality	44			
Prevalence of depression with no treatment, %	Good quality	55	0.05	0.00 (1.00)	+
	Poor quality	72			
Prevalence of bladder or bowel incontinence, low risk, %	Good quality	51	<0.001	0.59 (<0.001)	+
	Poor quality	92			
Prevalence of bladder or bowel incontinence, high risk, %	Good quality	53	<0.001	0.28 (<0.001)	+
	Poor quality	82			
Prevalence of indwelling catheters, low risk, %	Good quality	54	0.01	0.42 (<0.001)	+
	Poor quality	77			
Prevalence of indwelling catheters, high risk, %	Good quality	53	<0.001	0.39 (<0.001)	+
	Poor quality	83			
Prevalence of faecal impaction, %	Good quality	55	0.05	0.15 (0.12)	+
	Poor quality	72			
Prevalence of bedfast residents, %	Good quality	51	<0.001	0.66 (<0.001)	+
	Poor quality	92			
Lack of training/skill practice or range of motion for mobility dependent residents, %	Good quality	55	0.05	0.09 (0.34)	+
	Poor quality	72			
Prevalence of daily physical restraints, %	Good quality	53	<0.001	0.53 (<0.001)	+
	Poor quality	82			
Prevalence of little or no activity, %	Good quality	54	<0.001	0.36 (<0.001)	+
	Poor quality	79			
Prevalence of stage 1–4 pressure ulcers, low risk, %	Good quality	54	<0.001	0.31 (<0.001)	+
	Poor quality	78			
Proportion of registered nurses on the staff, %	Good quality	54	0.02	0.16 (0.09)	+
	Poor quality	75			
Proportion of single rooms, %	Good quality	56	0.10	0.36 (<0.001)	+
	Poor quality	71			
Prevalence of any injury, %	Good quality	57	0.95	−0.09 (0.36)	+/−
	Poor quality	58			
Incidence of new fractures, %	Good quality	58	0.81	−0.10 (0.20)	+/−
	Poor quality	56			
Use of nine or more different medications	Good quality	58	0.80	−0.06 (0.52)	+/−
	Poor quality	56			

continued

Table 3 *continued*

Variable	Quality groups	Mean rank in efficiency	Asymptotic sig. ¹	Correlation between efficiency scores and continuous quality indicators in pooled data (<i>n</i> = 114), <i>P</i> values in parentheses	Association between poor quality and high efficiency ²
Incidence of cognitive impairment, %	Good quality	58	0.75	0.02 (0.85)	+/-
	Poor quality	55			
Incidence of decline in late loss activities of daily living, low risk, %	Good quality	57	0.52	0.01 (0.91)	+/-
	Poor quality	62			
Incidence of decline in late loss activities of daily living, high risk, %	Good quality	57	0.86	0.01 (0.95)	+/-
	Poor quality	59			
Incidence of decline in late loss activities of daily living, %	Good quality	58	0.70	-0.15 (0.16)	+/-
	Poor quality	55			
Incidence of decline in range of motion, low risk, %	Good quality	58	0.54	-0.09 (0.32)	+/-
	Poor quality	53			
Incidence of decline in range of motion, high risk, %	Good quality	59	0.22	-0.19 (0.05)	+/-
	Poor quality	48			
Incidence of decline in range of motion, %	Good quality	59	0.16	-0.10 (0.29)	+/-
	Poor quality	46			
Prevalence of urinary tract infections, %	Good quality	58	0.63	0.03 (0.76)	+/-
	Poor quality	54			
Prevalence of weight loss, %	Good quality	56	0.28	0.04 (0.65)	+/-
	Poor quality	65			
Prevalence of antipsychotic use, in the absence of psychotic and related conditions, %	Good quality	56	0.28	0.11 (0.27)	+/-
	Poor quality	66			
Prevalence of anti-anxiety/hypnotic use, %	Good quality	59	0.24	-0.09 (0.35)	+/-
	Poor quality	49			
Prevalence of hypnotic use more than two times in last week, %	Good quality	57	0.50	0.16 (0.08)	+/-
	Poor quality	62			
Prevalence of stage 1–4 pressure ulcers, high risk, %	Good quality	58	0.57	0.06 (0.54)	+/-
	Poor quality	53			
Incidence of new pressure ulcers, %	Good quality	56	0.28	0.13 (0.18)	+/-
	Poor quality	65			
Prevalence of diagnosis or symptoms of depression, %	Good quality	60	0.11	-0.17 (0.07)	+/-
	Poor quality	46			
Prevalence of bladder or bowel incontinence, %	Good quality	58	0.13	0.09 (0.33)	+/-
	Poor quality	55			
Prevalence of occasional or frequent bladder or bowel incontinence without a toileting plan, %	Good quality	55	0.72	0.23 (0.01)	+/-
	Poor quality	68			
Prevalence of indwelling catheters, %	Good quality	56	0.33	0.28 (0.01)	+/-
	Poor quality	65			

¹Asymptotic significance (significant *P*-values) indicates statistical difference between the groups.²+ denotes that poor quality is associated with high efficiency, - denotes that poor quality is associated with low efficiency, and +/- denotes no clear positive or negative association between quality and efficiency.

Our results revealed a clear association between efficiency and quality in the prevalence-type of quality indicators in the time-consuming areas only. However, incidence-type indicators—such as the incidence of decline in the late loss of activities of daily living and the incidence of decline in the range of motion—were the most important indicators to

support our theory that caregivers adapt non-pro-active nursing patterns in the absence of adequate staffing ratios with regard to patients' needs. In the analysis, incidence-types of indicator, however, failed to show significant relationships, conceivably due in part to the small number of observations or a weak relationship.

Due to the cross-sectional nature of our study, the observed relationships are associative. The efficiency estimates would also have been more consistent in panel data and the longitudinal follow-up of patients might have revealed causal relationships. The clinical quality indicators used in this study may also be associated with each other. For instance, there may be a positive correlation between the prevalence of bedfast residents and pressure ulcers. Although our non-parametric approach enabled the use of all quality indicators in the analysis separately, one might consider using factor analysis for combining different quality indicators. However, for the time being in Finland these clinical quality indicators are routinely used for benchmarking purposes. Our results showed that efficiency was associated with certain types of quality indicator or quality dimension, which also favours the possibility of reducing the number of quality dimensions. The results also indicate that the selection of quality variables may affect the interference if certain quality dimensions are omitted. The strength of the non-parametric approach is that several quality indicators can be used in the analysis.

There seems to be clear improvement potential both in productive efficiency and quality of care, which may result in better outcomes and a better utilization of labour input and cost containment in the long run, although the population is ageing and probably more resources will be needed for the care of older persons. The results support the fact that efficiency and quality of care are essential aspects of management and performance measurement. If there is an intention to increase productive efficiency in elderly care—as there is in Finland—then quality has to be taken into account, for instance by using quality monitoring systems, although the empirical evidence on the association between quality and efficiency is still not fully consistent [11,12]. The association between efficiency and poor quality of care was detected in some areas associated with non-pro-active nursing practice and the consequences of such patterns. Further studies with a greater sample of wards and with longitudinal construction are warranted to show whether or not the results of this study indicate a true trade-off between productive efficiency and care.

Acknowledgements

This study has been funded by the Academy of Finland (Academy decision notification number 70034), the National Research and Development Centre for Welfare and Health (STAKES), and the Doctoral Programme in Social and Health Services Management and Economics of the University of Kuopio. Acknowledgement is also given to interRAI.

References

- Newhouse J. Toward a theory of nonprofit institutions: an economic model of a hospital. *Am Econ Rev* 1970; **60**: 64–74.
- Weech-Maldonado RJ, Neff G, Mor V. Does quality of care lead to better financial performance? The case of the nursing home industry. *Health Care Manage Rev* 2003; **28**: 201–216.
- Hirschman AO. *Exit, Voice, and Loyalty*. Cambridge, MA: Harvard University Press, 1970.
- Farrell MJ. The measurement of productive efficiency. *J R Stat Soc* 1957; **A120**: 253–281.
- Björkgren M, Häkkinen U, Linna M. A comparison of methods for determining efficiency of Finnish long-term care units. *Sos Laaketiet Aikak* 2000; **37**: 216–222 (in Finnish).
- Laine J, Linna M, Häkkinen U, Noro A. Measuring the productive efficiency and clinical quality of institutional long-term care for the elderly. *Health Econ* 2004; **14**(3): 245–256.
- Noro A, Finne-Soveri H, Björkgren M et al. *Introduction of the RAI-System and Experiences of Benchmarking of Long-term Care for the Elderly*. [RAI-tietojärjestelmän esittely ja kokemuksia pitkäaikaishoidon benchmarkingista.] RAI -raportti 1/2000. Helsinki: Stakes, 2001 (in Finnish).
- Berg K, Mor V, Morris J, Murphy KM, Moore T, Harris Y. Identification and evaluation of existing nursing homes quality indicators. *Health Care Financ Rev* 2002; **23**: 19–36.
- Kane R. Assuring quality in nursing home care. *J Am Geriatr Soc* 1998; **46**: 232–237.
- Dellefield ME. The relationship between nurse staffing in nursing homes and quality indicators. *J Gerontol Nurs* 2000; **27**: 14–28.
- Knox KJ, Blankmeyer EC, Stutzman JR. Organizational efficiency and quality in Texas nursing facilities. *Health Care Manag Sci* 2003; **6**: 175–188.
- Rosko M, Chilingirian J, Zinn J, Aaronson W. The effects of ownership, operating environment, and strategic choices on nursing home efficiency. *Med Care* 1995; **33**: 1001–1021.
- Chou S-Y. Asymmetric information, ownership and quality of care: an empirical analysis of nursing homes. *J Health Econ* 2002; **21**: 293–311.
- Morris JN, Hawes C, Fries BE et al. Designing the National Resident Assessment Instrument for nursing homes. *Gerontologist* 1990; **30**: 293–307.
- Fries BE, Schneider DP, Foley WJ, Gavazzi M, Burke R, Cornelius E. Refining a case-mix measure for nursing homes: Resource Utilization Groups (RUG-III). *Med Care* 1995; **32**: 668–685.
- Hawes C, Morris J, Phillips C, Fries B, Murphy C, Mor V. Development of the nursing home Resident Assessment Instrument in the USA. *Age Ageing* 1997; **26** (Suppl 2): 19–25.
- Zimmermann DR, Karon SL, Arling G et al. Development and testing of nursing home quality indicators. *Health Care Financ Rev* 1995; **16**: 107–140.
- Carpenter I, Perry M, Challis D, Hope K. Identification of registered nursing care of residents in English nursing homes using the Minimum Data Set Resident Assessment Instrument (MDS/RAI) and Resource Utilisation Groups version III (RUG-III). *Age Ageing* 2003; **32**: 7.
- Björkgren M, Häkkinen U, Finne-Soveri H, Fries B. Validity and reliability of Resource Utilization Groups (RUG-III) in Finnish long-term care facilities. *Scand J Public Health* 1999; **27**: 228–234.
- Carrillo E, Garcia-Altes A, Peiro S, Portella E. System for the classification of patients in mid and long-term care facilities:

- Resource Utilization Groups, version III. Validation in Spain. *Revista de Gerontologia* 1996; **6**: 276–284.
21. Charnes A, Cooper WW, Rhodes E. Measuring the efficiency of decision making units. *Eur J Oper Res* 1978; **2**: 429–444.
 22. Björkgren M, Häkkinen U, Linna M. Measuring efficiency of long-term care units in Finland. *Health Care Manag Sci* 2001; **4**: 193–200.
 23. Stuck AE, Walther J, Nikolaus T, Bula CJ, Hohmann C, Beck JC. Risk factors for functional status decline in community-living elderly people: a systematic literature review. *Soc Sci Med* 1999; **48**: 445–469.
 24. Castle NG. Nursing homes with persistent deficiency citations for physical restraint use. *Med Care* 2002; **40**: 868–878.
 25. van Marum RJ, Ooms ME, Ribbe MW, van Eijk JT. The Dutch pressure sore assessment score or the Norton scale for identifying at-risk nursing home patients? *Age Ageing* 2000; **29**: 63–68.
 26. Evans LK, Strumpf NE. Tying down the elderly: a review of the literature on physical restraint. *J Am Geriatr Soc* 1989; **37**: 65–74.
 27. Capezuti E, Strumpf NE, Evans LK, Grisso JA, Maislin G. The relationship between physical restraint removal and falls and injuries among nursing home residents. *J Gerontol A Biol Sci Med Sci* 1998; **53**: M47–52.
 28. Volicer L, Hurley AC. Management of behavioral symptoms in progressive degenerative dementias. *J Gerontol A Biol Sci Med Sci* 2003; **58**: M837–845.
 29. Martin H, Slyk MP, Deymann S, Cornacchione MJ. Safety profile assessment of risperidone and olanzapine in long-term care patients with dementia. *J Am Med Dir Assoc* 2003; **4**: 183–188.
 30. Vaarama M, Luomahaara J, Peiponen A, Voutilainen P. *The Whole Municipality Working Together for Older People*. Report. Helsinki: Stakes, 2001.

Accepted for publication 4 February 2005



Measuring the productive efficiency and clinical quality of institutional long-term care for the elderly

Juha Laine*, Miika Linna, Unto Häkkinen and Anja Noro

Centre for Health Economics at Stakes—CHESS, Finland

Summary

The authors consider the association between productive efficiency and clinical quality in institutional long-term care for the elderly. Cross-sectional data were collected from 122 wards in health-centre hospitals and residential homes in Finland in 2001. Productive efficiency was measured in terms of technical efficiency, which was defined as the unit's distance from the (best practice) production frontier. The analysis employed stochastic production frontier estimation, where technical inefficiency in the production function was specified to be a function of ward characteristics and clinical quality of care. Several quality indicators based on the Resident Assessment Instrument, such as prevalence of pressure ulcers and depression with no treatment, were used in the analysis. The results did not reveal systematic association between technical efficiency and clinical quality of care. However, the prevalence of pressure ulcers, indicating poor quality of care was associated with technical efficiency, a fact which highlights the importance of including quality measures in the assessment of efficiency in long-term care. Copyright © 2004 John Wiley & Sons, Ltd.

Keywords productive efficiency; clinical quality indicators; stochastic frontier; resident assessment instrument; long-term care

Introduction

In developed nations, the ageing of populations and the growing need for long-term care highlight the importance of economic evaluation of elderly care as a tool for policy-making and as a management instrument. In Finland, it has become apparent that the scarcity of qualified labour and limited budgets set a strict framework for providers and producers. Consequently, the allocation of resources and their effective use are crucial issues from the societal perspective. The success of this allocation depends on many factors, the most important of which are the actions taken by providers (strategic planning) and those taken

by producers (operative management). Unfortunately, the providers of health and social services in Finland lack appropriate measures of efficiency for strategic planning. Despite the common belief that there exists a trade-off between unit costs and quality, comprehensive quality measures are seldom available for strategic or operative reporting.

In the evaluation of productive efficiency and cost efficiency, most providers need to rely on very crude measures, e.g. per diem cost, as the basis for decisions. On the other hand, from the perspective of patients and their relatives the efficiency of services may be too abstract a concept; for them the most important and concrete criterion for

*Correspondence to: STAKES, Lintulahdenkuja 4, P.O. Box 220, FIN-00531 Helsinki, Finland. E-mail: juha.laine@stakes.fi

good service is usually experienced quality. Since efficiency and quality are fundamental but also potentially conflicting objectives, they should be simultaneously taken into account in economic evaluation and decision-making.

Technical efficiency is an important determinant of cost efficiency (unit costs). It has been estimated that approximately half of the differences in cost efficiency in Finnish long-term care can be attributed to technical efficiency [1]. In some empirical studies where cost functions were applied, the results have suggested a trade-off or association between costs and quality, but some studies indicate the opposite [2–6]. Still, efficiency and quality are most often considered individually in separate studies [7–10]. Hence, it is possible to argue that there is still a need for an integrative perspective which explores the relationship between the efficiency, costs and quality of care. The objective of this study was to evaluate the relationship between technical efficiency and clinical quality of care, and to understand the effects of ward characteristics and quality of care on technical efficiency.

The quantity–quality trade-off in Finland

In Finland, municipalities ($N = 446$) are responsible for providing the resident population with adequate social and health services. Financing is based on state subsidies, local taxes and user fees. Institutional elderly care is provided mainly in public residential homes and health-centre hospitals. For the time being, the growing number of private service producers are concentrating more on the provision of care at home and sheltered housing [11]. One of the main problems in elderly care is that public long-term care facilities lack (properly trained) staff. However, more resources will need to be allocated to long-term care in the future, because the population is ageing rapidly. Public institutions and their management are having to perform a balancing act between cost minimisation, increasing demand and adequate care. It is assumed in this study that facilities and wards try to minimise input (staff and capital) while holding output (inpatient days) constant or to maximise output while holding input fixed. In this context a trade-off between quantity (and costs) and quality of care may occur.

In Finland, public discussion concerning problems in the quality of long-term care for elderly

has resembled the discussion in the US in the 1980s [12,13]. Despite widespread public concern, nationwide actions aimed at ensuring and monitoring quality have concerned only staffing, and not other quality dimensions. Comprehensive quality standards or monitoring are also needed, because the possibility for patients to evaluate and demand quality is restricted, especially if cognition is impaired or if relatives are absent.

The relationship between quantity and quality is an important aspect in the Finnish, i.e. in the public financed and organised health care system. An individual's opportunity to choose a producer is restricted in practice. A person's long-term care institution is usually determined by the area where he or she lives. Furthermore, the link between inpatient fees and quality of care can be quite blurred. The progressive inpatient day fee is determined in advance by the management of the facilities, in line with budget constraints. The number of private producers is also limited, and their services are available mostly for wealthy people in the largest cities. In the context of asymmetric information and restricted choice, particular attention should be paid to the association between quantity and quality of the care.

The most common criticism levelled at efficiency comparisons concerns the measurement bias resulting from the omission of quality information [14]. According to this view [15], quantity and quality are considered to be separate products and different dimensions in the production possibility set. If an increase in patient days (output) or a decrease in costs (inputs) is achieved by debasing the quality of care, there exists a cost–quality trade-off in long-term care. In that case, the best practice production frontier depends on the quality level, and the production units may compensate for lower efficiency through higher quality and vice versa. This trade-off may occur if quality is measured as the number of private rooms in a ward or of staff per resident, because these factors increase the usage of labour and capital. If quality is instead measured in terms of clinical measures or client satisfaction, the trade-off is not so evident. For instance, actions aimed at achieving better client satisfaction, such as better communication with patients and relatives, do not necessarily increase costs. If it is possible and meaningful to make clear trade-offs between quantity and quality in the production process, then there is certainly a need for quality-controlled efficiency evaluations.

Definitions of quality

In the theoretical input–output model of production, quality can be defined as something that is characterised by (i) the type and essence of outputs or outcomes, (ii) the type and essence of inputs used or (iii) certain events in the transformation process which can not be clearly defined as inputs or outputs ('side-effects' occurring simultaneously in the process). This typology could be explored alongside the common empirical methods in quality measurement. The three main approaches to the measurement of quality of care in institutional long-term care for the elderly have traditionally been (i) the collection of standardised measures of output, processes and outcomes of care, (ii) a survey of patient or relative satisfaction or (iii) a survey of facilities' capability to supply care (staffing, physical standards) [16–18]. In practice, a ward in a residential home or in a health-centre hospital may perform well in some dimension of quality, but poorly in other dimensions [19,20], a fact which may complicate the application of quality measurement in empirical studies on efficiency.

Output quality

Incorporating quality in the output specification is intuitively appealing, and examples of this approach can be found [4,21,22]. In many industrial or even service branches, quality may be regarded as something that classifies the product as belonging to a certain product category. Products can intentionally be manufactured at many quality levels or standards, and the acceptance of the product is left to the market. In the traditional input–output model of production, the various quality levels then span a multidimensional output space. As an example, let us assume that two types of output (y_1, y_2) can be produced at two levels of quality: high and low. In this case, the output can be represented using a four-dimensional output vector (y_1 -high, y_1 -low, y_2 -high, y_2 -low). However, the output in institutional elderly care (number of inpatient days) should be quite homogenous. In long-term care, it should be unacceptable to intentionally vary total quality (for instance, to explicitly produce 50% of inpatient days at poor quality) or to debase quality by increasing quantity—at least, not when quality

is characterised by outputs or outcomes of care. In this study we do not incorporate quality into case-mix weighted inpatient days (output), because there is no extensive and composite quality measure for weighting the single output.

Input and structure quality

In a broad sense it may sometimes be justified to assess quality in terms of the input types and quantities. The resources allotted to each production unit vary, and it may well be that resources given to a ward or a facility set a framework within which there are (different) opportunities to produce good care, despite the fact that the nursing staff surely do the best they can. For instance, the use of different staff categories may have an influence on some dimensions of quality [23]. However, increasing the use of resources (e.g. money, person years) does not necessarily produce better outcomes (health and quality of life) or quality of care, especially in the presence of inefficiency. The causality existing between resources and quality of care is quite indefinite.

Process and outcome quality

In institutional settings for elderly care, the quality of life of patients is partly dependent, for instance, on the avoidance of accidents and on delaying physical and mental degeneration. Adverse outcomes and adverse care processes such as falls and declines in activities of daily living (ADL) are captured by clinical measures. In this study we pay attention to clinical (objective) indicators that describe several clinical dimensions of quality of care [24]. Quality is not considered as an output or as types of input in the analysis, but rather as a factor in the production function and care process which may affect the technical efficiency of an unit. One dimension of outcome is also subjective measures of quality experienced by patients and relatives or perceived quality of employers, but these measures will be tested later on.

Data

Following Björkgren *et al.* [1], the analysis was based on ward-level measurement, which can be shown to be better than facility-level analysis for

many reasons. First of all, it is clearly preferable to use a smaller amount of aggregate data than would be necessary in facility-level analysis. Secondly, wards are units, to which resources such as staff are allocated by management. Moreover, the decisions concerning care, which are most crucial from the point of view of the quality and effectiveness of care, are also made at the ward level.

In this study we utilised the Resident Assessment Instrument (RAI), which is a tool for improving nursing and quality of care. The RAI for the nursing homes is based on the Minimum Data Set 2.0 (MDS 2.0) questionnaire, which is used by the nursing staff to assess patients in the ward [12]. In Finland each patient is assessed at admission and every six months or if the patient's condition changes distinctly. The MDS 2.0 is composed of over 300 variables and 18 core assessment items, including the patient's background, cognitive patterns and health conditions. The contents of the MDS enable one to construct many indices and indicators, such as the case-mix index (Resource Utilisation Groups), activities of daily living (ADL) and clinical quality indicators indicating adverse outcomes and adverse processes of care [25–29].

The data were collected in collaboration with participants in the 'Benchmarking and Implementation of RAI in Elderly Care in Finland' project [20]. The complete data were obtained from 122 wards in 39 facilities for the year 2001. Information on annual costs and full-time equivalent person-years for each ward was obtained from the finance offices of the facilities. Detailed instructions were given for assessing cost collection and the allocation of overheads. The information concerning staffing, unit type and occupancy rate (ward characteristics) was collected from the head nurses by means of a questionnaire. The case-mix index, patients' ages and quality indicators were obtained from our RAI database. While quality measures, patients' ages and case-mix index were originally measured at patient level, they were aggregated to the ward level for the analysis.

The case-mix adjustment in this study was applied to the output variable, which is the number of inpatient days. We calculated the patient days assigned to the RUG-III (Resource Utilisation Groups) classification for each MDS assessment per episode of care in the ward, and summed up the total for each patient. In Finland,

22-classification of RUG-III is used, and the cost weights for different patient groups vary from 0.48 to 2.52 [30]. The following formula was used for each patient in the ward during year 2001:

$$Y_k = \sum d_i w_i \quad (1)$$

where d_i is the patient days in RUG-III group, and w_i the cost weight in RUG-III group.

The RAI enables the use of over 20 clinical quality indicators [24], which were also available from the RAI database. In the first stage, these clinical indicators were considered from the contentual viewpoint by experts in nursing, who ascertained how they could be associated with quality of care in residential homes and health-centre hospitals, and whether they had been used in other studies as quality measures. Some RAI quality indicators were not suitable for this study from the beginning; this was so if (i) the indicator was an incidence measure and the number of observations from which it was calculated was too low, (ii) there was not enough variation in the variable, (iii) the quality indicator was not suitable for both health-centre hospitals and residential homes or (iv) the indicator was strongly correlated with other variables. In addition, the number of the observations ($n = 122$) limited the number of the indicators of inefficiency effects. On the basis of the reasoning and criteria presented above, three quality indicators indicating adverse care processes and outcomes were chosen for further analysis. The indicators were prevalence (%) of weekly use of depressants and hypnotics, prevalence of depression with no treatment and prevalence of pressure ulcers. The prevalence figures were the mean values of measurements made during three RAI assessment periods (cross-sections) falling between 16 September 2000 and 15 May 2002. They were calculated from several cross-sections in order to obtain more consistent quality measures and to diminish measurement errors and randomness. The descriptive statistics for the final data are given in Table 1.

When quality of care is measured as the prevalence of adverse outcomes or adverse care processes, we also have to take into account the fact that both very low and very high prevalence may sometimes indicate poor quality. If there is not enough staff and supervision, patients may fall easily. On the other hand, patients may be encouraged to move to maintain their physical

Table 1. Descriptive statistics of the variables

Variable name	Description	Mean (S.D.)	Source (year 2001)
<i>Output</i>			
Case-mix weighted patient days	Calculation of patient days associated with the RUG-III classification for each MDS assessment within an episode of care (= ward's total case-mix weighted patient days)	10 824 (4706)	RAI database/MDS
<i>Inputs</i>			
Full-time equivalent person years for registered nurses (RN)	Full time equivalent person years	3.9 (2.7)	Finance office
Full-time equivalent person years for licensed practical nurses (LPN)	Full time equivalent person years	10 (5.1)	Finance office
Full-time equivalent person years for aides (AIDES)	Full time equivalent person years	2.4 (1.9)	Finance office
Unit size	Number of beds	30.6 (12.4)	Questionnaire
<i>Ward characteristics</i>			
Facility type	0 = Residential home 1 = Health-centre hospital	—	RAI database/MDS
Occupancy rate	Occupancy rate	97.5 (11.4)	Finance office
Ward's specialisation	0 = Mixed patients 1 = Psycho-geriatric patients, i.e. at least 60% of patients belonging to RUG-III/22 groups IA, IB, PC, PD, BA, BB	—	RAI database/MDS
Mean age	Patients' average age	81.3 (8.7)	RAI database/MDS
<i>Clinical quality indicators^a</i>			
Prevalence of (1–4 stage) pressure ulcers	Continuous (%) outcome measure	11.0 (7.0)	RAI database/MDS
Prevalence of weekly use, depressants and hypnotics	Continuous (%) process measure	60.0 (14.0)	RAI database/MDS
Prevalence of depression with no treatment	Continuous (%) process/outcome measure	14.0 (7.0)	RAI database/MDS

^aHigher prevalence (%) indicates poorer quality.

condition. However, if restraints are used and patients are bound to beds and chairs, there is actually no possibility of falling, so that the prevalence of falls will probably be low. For this reason some clinical quality measures are more

continuous by nature (i.e. the higher the prevalence, the worse the quality) and some are more dichotomous (i.e. low and high prevalences indicate poor quality). This is illustrated in Figure 1 in terms of pressure ulcers and falls.

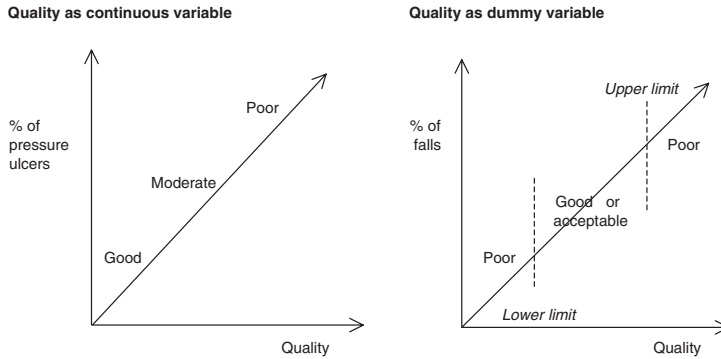


Figure 1. Quality specified as continuous and as dummy variable

The recommended lower and upper limits for good quality have not yet been defined in Finland as in the USA [31], which is why we also constructed a dummy quality variable based on outlier observations. Dummy is 0 if a ward's score falls between 10% and 90% in the distribution for all wards, and 1 if the score is in the lowest or highest decile. We also tried the lower and upper limits of 5% and 95%, respectively, in the analysis. The US limits (thresholds) for good quality were also tested, but because of the differences in the client structures and nursing cultures, some quality thresholds might not be transferable to Finland. In addition, squared variable terms were also used to test the U-shaped effect.

Methods

Optimising the use of labour and capital in long-term care assumes information and knowledge concerning production processes. This can be defective among managers, and facilities often fail to achieve efficient use of resources [10]. Inefficiency in production can be modelled by employing methods for calculating efficiency. The main approaches for calculating efficiency and productivity are the parametric and non-parametric methods. The parametric approach is based on the econometric tradition of cost and production functions [32,33] while the non-parametric method, data envelopment analysis (DEA), is based on linear programming [34,35]. The aim of most empirical studies is not only to find the most efficient decision-making unit(s), but also to find

the determinants of inefficiency. Two approaches have traditionally been employed in the parametric estimation of inefficiency: 'single-stage estimation' and 'two-stage estimation'.

The two-stage approach has been widely used in previous efficiency studies on elderly care [7,10,36]. In the first stage, efficiency scores are calculated using stochastic frontier estimation or DEA. In the second stage, regression analysis is used in order to explain the efficiency scores as a function of ward characteristics and other factors. However, it has been argued that the parameter estimates in second-stage regression are non-consistent, because the technical inefficiency obtained from the first stage is not stochastic [37,38]. It might be a function of some firm-specific factors, which may then be used as explanatory variables in the second-stage regression. It has been suggested that this pitfall of the two-stage approach can be avoided by applying single-stage estimation (inefficiency effect estimation), in which the production or cost frontier and inefficiency effects are estimated simultaneously [39,40]. The inefficiency error component is a function of various firm- and environment-specific factors and a random component. The method is based on maximum likelihood estimation. It is used in this study in order to evaluate if technical inefficiency is an explicit function of certain firm-specific factors (ward characteristics and quality of care).

The general (linear) formulation of the stochastic production frontier model is [41]

$$y_i = \alpha + \sum_j \beta_j x_{ij} + v_i - u_i \quad (2)$$

where y_i is the output, x_j are input variables, v_i is a random-error term and u_i is the non-negative inefficiency component. v and u are distributed as follows:

$$v_i \sim N[0, \sigma_v^2]$$

$$u_i = |U|, U \sim N[0, \sigma_u^2]$$

In technical inefficiency effect estimation, u is treated as a function of some firm- and environment-specific factors. The general formulation of the technical inefficiency effect model is

$$y_i = \alpha + \sum_j \beta_j x_j + v_i - u_i \quad (3)$$

where β_j, x_j, v_i are as in (2). u_i is composed of non-negative random variables that are assumed to be independently distributed as truncations of the $N(\delta Z_i, \sigma_u^2)$ distribution. Inefficiency effects, denoted Z_i , arise by truncation at zero of the normal distribution with mean μ_i and variance σ_u^2 . δ_i -coefficients are unknown parameters to be estimated together with variance parameters $\sigma_s^2 = \sigma_v^2 + \sigma_u^2$ and $\gamma = \sigma_u^2 / \sigma_s^2$. In this study the model for the technical inefficiency effects in the stochastic production frontier equation was defined by

$$\begin{aligned} \mu_i = & \delta_0 + \delta_1(\text{UNIT}_i) + \delta_2(\text{UNIT2}_i) + \delta_3(\text{WARD}_i) \\ & + \delta_4(\text{FACILITY}_i) + \delta_5(\text{OR}_i) + \delta_6(\text{AGE}_i) \\ & + \delta_7(\text{UNITOR}_i) + \delta_8(\text{WARDUNIT}_i) \\ & + \delta_9(\text{MED}_i) + \delta_{10}(\text{PRESS}_i) \\ & + \delta_{11}(\text{DEPRES}_i) + w_i \end{aligned} \quad (4)$$

where UNIT is unit size, UNIT2 is unit size squared, WARD is ward's specialisation, FACILITY is facility type, OR is occupancy rate, AGE is mean age of patients, UNITOR is interaction between unit size and occupancy rate, WARDUNIT is interaction between ward specialisation and unit size, MED is the weekly use of depressants and hypnotics, PRESS is the prevalence of pressure ulcers and DEPRESS is the prevalence of depression without treatment. w_i is defined by the truncation of the normal distribution with zero mean and variance, σ^2 , such that u_i is non-negative.

A substantial number of variables—such as number of purchased services, team/module work, patient turnover, long-term or short-term ward, job rotation and scope—were omitted from the final frontier model because they were

not systematically significant in the preliminary models.

The analysis proceeded as follows: In the first stage the likelihood ratio test, which was calculated as $\lambda = -2[\ln(\text{LH}_0)] - [\ln(\text{LH}_1)]$, rejected linear and logarithmic model specification of the stochastic frontier production function. Thus we applied the Box–Cox transformation to the output and inputs in the production function using LIMDEP [41]. The Box–Cox transformation is $y^{(\theta)} = (y^\theta - 1)/\theta$. The more flexible functional forms such as translog were not used, because the number of the observations (and degrees of freedom) was limited. We then estimated the production function and inefficiency effects simultaneously using FRONTIER 4.1c [42] software and specifications proposed by Battese and Coelli [39,42]. The programme uses a three-step procedure in estimating the maximum likelihood estimates of the parameters of a stochastic frontier production function and inefficiency effects.

Results

Since there was no preliminary knowledge of the model specification, several models were tested to estimate the stochastic production function with the technical inefficiency component. A number of restrictions were tested in our inefficiency effect model. Table 2 presents the tested hypotheses and results which were obtained using the likelihood-ratio test. All three of the following null hypotheses were rejected at the 0.01 level; (i) the inefficiency effects are absent from the model, (ii) the inefficiency effects as a group do not predict efficiency, (iii) the inefficiency effects do not have an intercept parameter. The correct critical values for the hypothesis involving $\gamma = 0$ is taken from Table 1 in Kodde and Palm [43].

The stochastic frontier production function and inefficiency estimates are presented in Table 3. Three quality measures and eight ward characteristics were used as inefficiency effects. The parameters of the input variables in the production function were all positive, except for the full-time equivalent person-years for aides. A positive value of an inefficiency effect (delta parameters $\delta_0 - \delta_{11}$) implies a positive relationship with inefficiency. The prevalence of pressure ulcers (Q2) was the only quality indicator which was significantly associated with technical efficiency (i.e. negatively

Table 2. Tests of hypotheses for parameters of the inefficiency stochastic production function

Null hypothesis	χ^2 -statistic	$\chi^2_{0.95}$ -value	Decision
Stochastic frontier			
$H_0: \gamma = \delta_0 = \dots = \delta_{11} = 0$	94.54	20.41	Reject H_0
$H_0: \delta_1 = \dots = \delta_{11} = 0$	89.83	18.55	Reject H_0
$H_0: \delta_0 = 0$	8.06	3.84	Reject H_0

Table 3. Maximum likelihood estimates for parameters of the stochastic frontier with inefficiency-effects model involving ward-specific variables

Variable	Parameter	Estimates ^a
<i>Stochastic frontier^b</i>		
Constant	β_0	0.082 (0.299)
Box–Cox transformed FTP for registered nurses ^c	β_1	0.066 (2.797)
Box–Cox transformed FTP for licensed practical nurses	β_2	0.055 (0.994)
Box–Cox transformed FTP for aides	β_3	−0.000 (−0.029)
Box–Cox transformed unit size (number of beds)	β_4	0.471 (8.923)
<i>Inefficiency effects</i>		
Constant	δ_0	4.772 (3.429)
Unit size	δ_1	−0.047 (−1.351)
Unit size square	δ_2	−0.000 (−0.595)
Ward's specialisation ^c	δ_3	0.557 (2.153)
Facility type ^c	δ_4	−0.573 (−3.458)
Occupancy rate	δ_5	−0.052 (−4.004)
Mean age	δ_6	−0.003 (−0.572)
Unit size × occupancy rate	δ_7	0.001 (2.951)
Special ward × unit size	δ_8	−0.009 (−1.191)
Prevalence of weekly use of depressants and hypnotics	δ_9	0.005 (1.332)
Prevalence of pressure ulcers	δ_{10}	−0.026 (−2.496)
Prevalence of depression with no treatment	δ_{11}	0.009 (1.445)
<i>Variance parameters</i>		
	σ_s^2	0.126 (4.851)
	γ	0.979 (29.149)
Likelihood function		−3.983

^aAsymptotic *t*-test in parenthesis.
^bBox–Cox transformation is $\theta = 0.21707$ and $\lambda = 0.38041$.
^cFTP, Full time equivalent person years.
Ward's specialisation: 0 = mixed patients, 1 = psycho-geriatric patients.
Facility type: 0 = residential home, 1 = health-centre hospital.

associated with inefficiency in the model). Higher prevalence (lower quality) increased technical efficiency, suggesting an association between technical efficiency and clinical quality. However, the prevalence of depression with no treatment (Q3) and the prevalence of weekly use of depressants and hypnotics (Q1) were negatively—though not significantly—associated with efficiency. This result is in contradiction with the case of pressure

ulcers, indicating that low quality may also be associated with technical inefficiency. The signs of different quality variables were not parallel in the model, and there is thus no fully consistent evidence that clinical quality in general is a significant determinant of technical inefficiency. Using the parameter estimates of model described in Table 3, it was possible to assess more explicitly the impact of each of the quality

variables on efficiency. Holding other variables constant, we calculated the change in efficiency units when an average ward moves from the lowest decile or quartile to the highest decile or quartile in the quality distributions. Calculations were performed separately for each of the quality indicators tested. In Figure 2, a higher value (%) in the quality indicators means lower quality. A shift from the lowest (good quality) to the highest (poor quality) group in the case of variables Q1 and Q3 seemed to have only a moderate effect on technical efficiency. A shift from the first quartile to the third quartile to the third quartile would mean a change of between 0.04 and 0.05 efficiency units for an average ward. Since the average level of technical efficiency was 0.84, a change of 0.05 efficiency units corresponds to a 6% change in average efficiency. The effects of quality variable Q2 were marked and inverse. Moving from the first quartile to the third quartile was found to result in an increase of 0.11 inefficiency units. When the upper and lower deciles of the distribution were used, the impact was as large as 0.24 inefficiency units.

It is possible that quality has different effects in different types of wards and facilities, and this may also reflect differences in case-mix. Hence, we also tested all interactions between facility type/ward

type and dummy/continuous quality indicators, but these proved not to be significant. Furthermore, changing continuous quality variables to dummy or squared variables did not affect the results.

In Table 3, ward characteristics such as location in health-centre hospitals and occupancy rate were significantly and negatively associated with technical inefficiency, while psycho-geriatric wards, which mostly take care of patients with dementia, tended to exhibit lower technical efficiency. We also tested several interactions between ward characteristics; only the interaction between unit size and occupancy rate was found to be significant and positively associated with technical inefficiency, although the effect on efficiency was rather weak. This suggests that increasing the number of beds and the occupancy rate simultaneously does not improve technical efficiency.

Discussion

In this study we applied the technical inefficiency effect estimation method proposed by Battese and Coelli to explore the effect of selected quality

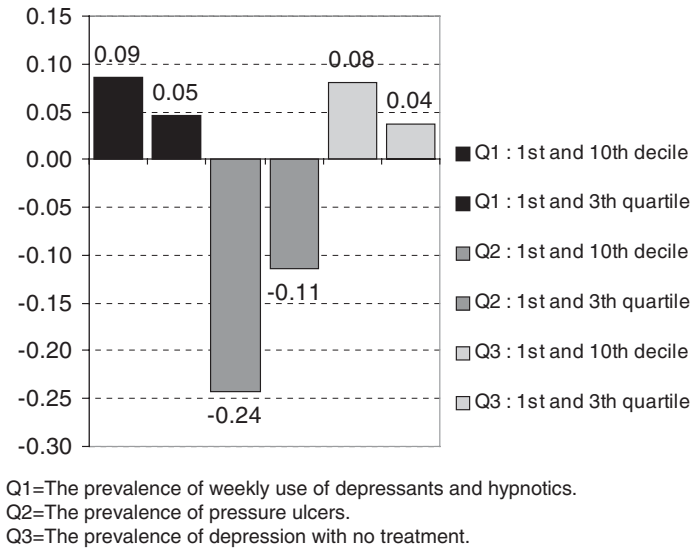


Figure 2. The effect on inefficiency of transition from the lowest to the highest (poorest) deciles and quartiles in the quality distribution

measures on efficiency [39]. Three continuous quality variables were chosen on the basis of expert reasoning and statistical criteria (e.g. statistical significance, multicollinearity). These clinical quality indicators reflecting outcome and process quality were not systematically associated with technical inefficiency.

Although low quality measured in terms of the prevalence of pressure ulcers was associated with efficiency, the results did not consistently support the assumption that productive efficiency compromises the total clinical quality of care. Quality measures based on depression medication and treatment were in fact positively correlated with efficiency. Thus units may attain high scores in some dimensions of clinical quality and still reach high levels of technical efficiency.

However, the prevalence of pressure ulcers in a ward seemed to have a effect on technical efficiency. This effect was demonstrated by calculating the change in the predicted efficiency score of a ward moving from the lowest (best quality) decile/quartile to the highest (poorest quality) decile/quartile. A transition from the lowest to the highest decile was found to increase efficiency by as much as 0.24 efficiency units in an average ward; thus by decreasing quality, higher efficiency is achieved. This effect was calculated by using the parameter estimates from the multivariate model, and thus the estimate is prone to statistical uncertainty. However, the result suggests that the prevalence of pressure ulcers is an important factor in efficiency comparisons and performance measurements.

Strategic and operative management in long-term care institutions should for many reasons pay particular attention to the quality of care, especially to the prevalence of pressure ulcers. First of all, there exists an economic association between quality and efficiency, and this affects staff sizing requirements and the price per inpatient day. Secondly, the unit prices are not comparable across facilities in the absence of quality information. Pressure ulcers, as a widely used quality indicator, therefore constitute an important top-level performance indicator and should be included in routine reporting to management. There is also a growing need for defining a tolerance level (thresholds) of the clinical quality indicators currently used in Finland.

Wards are expected to produce high or at least sufficient levels of quality. However, the results of the present study suggest that attaining these

standards may result in substantial costs and efficiency losses for many units. The information provided by our production model can be used to estimate the economic consequences of quality management programs. In the setting of strategic goals or benchmarks, management should take into account the observed link between quality and resource allocation, which can be used in the practical implementation of quality objectives.

The impact of the other quality variables—the weekly use of depressants and hypnotics, and depression with no treatment—may be considerably harder to explain. Institutional elderly care wards can sustain sufficient quality even at the highest levels of productive efficiency, at least when quality is defined by several clinical indicators. A possible explanation is that in the case of wards and facilities where the management and care staff are both interested in improving efficiency, they are also willing to control and improve quality by establishing quality programs. Furthermore, the working ethics and high professional skills of the nurses may prevent the occurrence and acceptance inactivating (passive) nursing practice, which cause adverse clinical outcomes and processes. In addition, it may be easier to control, monitor and improve clinical quality than is the case with other dimensions of quality. It has also been argued [16] that there is no composite score for evaluating overall quality. A facility may perform well in several areas but usually does not perform well in all. There is dispersion between and within different measures of quality. Because there are a number of dimensions to quality, no single process, outcome or other measure of quality should be used alone in the ranking of units. At some later time, it would be interesting to continue the analysis in order to see if the views of patients, relatives and staff members concerning quality are associated with efficiency, or whether the non-parametric method (DEA) would reveal more clearly the best cost or production function in the multidimensional quality space.

The average technical inefficiency among the wards was approximately 16%, which is close to what has been found earlier in long-term care. The ward characteristics associated with technical efficiency were roughly similar to those found in earlier studies [10]. Earlier comparative studies have revealed that data envelopment analysis and stochastic frontier analysis produce very similar technical efficiency estimates [10,44]. However,

these studies stress the importance of sensitivity analysis concerning the model structure [45]. In this study the analysis revealed that parameter estimates in the models were mildly sensitive to the model structure. One weakness of the study was that only cross-sectional data were available. While frontier estimation is possible with cross-sectional data, the unit-specific inefficiency scores would be more consistent with panel data [46]. However, the descriptive analysis of our previous RAI-project data (1999–2001) revealed that the variation in technical efficiency has been stable over the years.

The case-mix was taken into account in the output by using RUG-III weighting in inpatient days. Health centre-hospitals were found to be more efficient, even when patients' average age and case-mix were adjusted for. This may indicate that RUG-III case-mix adjustment does produce too low a cost-weight for certain patient groups—for instance, in the case of dementia patients with good physical functioning. It may be that some wards in residential homes may not receive enough weighting in inpatient days. Therefore the revision of RUG-III cost weights in Finland is important. However, this is a generally recognised problem in the RUG-III system. It is possible that the prevalence of pressure ulcers in a ward also reflects case-mix. We tried to take this into account by including the facility and ward types, interaction terms and case-mix variables in our model.

In the analysis, we tested several ward characteristics and clinical quality indicators, some of which have not been tested in previous studies. Our results support mild association between technical efficiency and some dimension of the clinical quality of care, but there still remains variation in technical efficiency, which was not explained by the set of exogenous factors, i.e. ward characteristics, or by the set of clinical quality indicators used in this analysis.

Acknowledgements

This study has been funded by the Academy of Finland, the National Research and Development Centre for Welfare and Health (Stakes) and the Doctoral Programme in Social and Health Services Management and Economics of the University of Kuopio. We thank Professor Hannu Valtonen for providing theoretical remarks and Professor Tim Coelli for the Frontier 4.1c software. The authors also express their gratitude to interRAI.

References

1. Björkgren M, Häkinen U, Linna M. Measuring efficiency of long-term care units in Finland. *Health Care Manage Sci* 2001; **4**: 193–200.
2. Carey K, Burgess JFJ. On measuring the hospital cost/quality trade-off. *Health Econ* 1999; **8**(6): 509–520.
3. Rosko M, Chilingirian J, Zinn J, Aaronson W. The effects of ownership, operating environment and strategic choices on nursing home efficiency. *Med Care* 1995; **33**: 1001–1021.
4. Blank JLT, Eggink E. A quality-adjusted cost function in a regulated industry: the case of Dutch nursing homes. *Health Care Manage Sci* 2001; **4**: 201–211.
5. Morey RC, Fine DJ, Loree SW, Retzlaff-Roberts DL, Tsubakitani S. The trade-off between hospital cost and quality of care. An explanatory empirical analysis. *Med Care* 1992; **30**: 677–698.
6. Waters T, Almagor O, Finkel S, Harter K, Bartolozzi P, Mitzen P, Lyons J, Edelman P, Fulton B, Hughes S. Understanding costs of home and community based services. *Manage Care Quality* 2001; **9**: 45–53.
7. Vitaliano D, Toren M. Cost and efficiency in nursing homes: a stochastic frontier approach. *J Health Econ* 1994; **13**: 281–300.
8. Rosko MD. Cost efficiency of US hospitals: a stochastic frontier approach. *Health Econ* 2001; **10**: 539–551.
9. Karon S, Zimmermann D. Using indicators to structure quality improvement initiatives in long-term care. *Quality Manage Health Care* 1996; **4**: 54–66.
10. Björkgren M, Häkinen U, Linna M. A comparative application of methods for determining efficiency of Finnish long-term care units. *J Social Med* 2000; **37**: 216–222.
11. SVT. Care and services for older people 2002. In *Finnish Official Statistics*. Social Security: I. Helsinki, 2003.
12. Morris JN, Hawes C, Fries BE, Phillips CD, Mor V, Katz S, Murphy K, Drugovich ML, Friedlob AS. Designing the national resident assessment instrument for nursing homes. *Gerontologist* 1990; **30**: 293–307.
13. Vaarama M, Luomahaara J, Peiponen A, Voutilainen P. *Koko kunta ikääntyneiden asialle. Näkökulmia ikääntyneiden itsenäisen selviytymisen sekä hoidon ja palvelun kehittämiseen*. Stakes: Helsinki, 2001.
14. Newhouse JP. Frontier estimation: how useful a tool for health economics? *J Health Econ* 1994; **13**: 317–322.
15. Newhouse JP. Toward a theory of nonprofit institutions: an economic model of a hospitals. *Am Econ Rev* 1970; **60**: 64–74.

16. Ikegami N, Hirdes JP, Carpenter I. Measuring the quality of long-term care in institutional and community setting. In *Measuring UP: Improving Health System Performance in OECD Countries*. OECD, Paris, 2002.
17. Seidman JJ, Greenfield S. Quality for good measure. *Sem Med Practice* 1998; **1**: 7–10.
18. Ovretveit J. *Evaluating Health Interventions*. Open University Press: Buckingham, 1998.
19. DesHarnais S, McMahon LFJ, Wroblewski R. Measuring outcomes of hospital care using multiple risk-adjusted indexes. *Health Services Res* 1991; **26**: 425–445.
20. Noro A, Finne-Soveri H, Björkrgren M, Häkkinen U, Laine J, Vähäkangas P, Kerppilä S, Storbacka R. *RAI-tietojärjestelmän esittely ja kokemuksia pitkäaikaishoidon benchmarkingista. RAI-raportti 1/2000*. Stakes: Helsinki, 2001.
21. Färe R, Grosskopf S, Roos P. Productivity and quality changes in Swedish pharmacies. *Int J Production Econ* 1995; **39**: 137–147.
22. Olesen O, Petersen N. Incorporating quality into data envelopment analysis: a stochastic dominance approach. *Int J Production Econ* 1995; **39**: 117–135.
23. Needleman J, Buerhaus P, Mattke S, Stewart M, Zelevinsky K. Nurse-staffing levels and the quality of care in hospitals. *New England J Med* 2002; **346**: 1715–1722.
24. Zimmermann DR, Karon SL, Arling G *et al.* Development and testing of nursing home quality indicators. *Health Care Finance Rev* 1995; **16**: 107–140.
25. Fries BE, Cooney Jr LM. Resource utilization groups. A patient classification system for long-term care. *Med Care* 1985; **23**: 110–122.
26. Fries BE, Schneider DP, Foley WJ, Gavazzi M, Burke R, Cornelius E. Refining a case-mix measure for nursing homes: resource utilization groups (RUG-III). *Med Care* 1995; **32**: 668–685.
27. Hawes C, Morris J, Phillips C, Mor V, Fries B, Nonemaker S. Reliability estimates for the minimum data set for nursing home resident assessment and care screening. *Gerontologist* 1995; **35**: 172–178.
28. Hawes C, Morris J, Phillips C, Fries B, Murphy C, Mor V. Development of the nursing home resident assessment instrument in the USA. *Age and Ageing* 1997; **26**(S2): 19–25.
29. Murphy C, Morris J. The resident assessment instrument: implications for quality, reimbursement and research. *Generations* 1995; **19**: 43–46.
30. Björkrgren M, Häkkinen U, Finne-Soveri H, Fries B. Validity and reliability of resource utilization groups (RUG-III) in Finnish long-term care facilities. *Scand J Public Health* 1999; **27**(3): 228–234.
31. Rantz MJ, Petroski GF, Madsen RW *et al.* Setting thresholds for MDS quality indicators for nursing home quality improvement reports. *J Quality Improvement* 1997; **23**: 602–611.
32. Aigner DJ, Lovell CAK, Schmidt P. Formulation and estimation of stochastic frontier production function models. *J Econometrics* 1977; **6**(2): 21–37.
33. Meeusen W, van den Broeck J. Efficiency estimation from Cobb–Douglas production functions with composed error. *Int Econ Rev* 1977; **18**: 435–444.
34. Farrell MJ. The measurement of productive efficiency. *J R Statist Soc A* 1957; **120**: 253–281.
35. Charnes A, Cooper WW, Rhodes E. Measuring the efficiency of decision making units. *Eur J Oper Res* 1978; 429–444.
36. Kooreman P. Nursing home care in The Netherlands: a nonparametric efficiency analysis. *J Health Econ* 1994; **13**: 301–216.
37. Kumbhakar S, Ghosh S, McGuckin J. A generalized production frontier approach for estimating determinants of inefficiency in US dairy farms. *J Business Econ Statist* 1991; **9**: 279–286.
38. Reifschneider D, Stevenson R. Systematic departures from the frontier: a framework for the analysis of firm inefficiency. *Int Econ Rev* 1991; **32**: 715–723.
39. Battese GE, Coelli TJ. A stochastic frontier production function incorporating a model for technical inefficiency effects. *Working Papers in Econometrics and Applied Statistics* 1993.
40. Battese GE, Coelli TJ. A model for technical inefficiency effects in stochastic frontier production function for panel data. *Empirical Econ* 1995; **20**: 325–332.
41. Greene W. *LIMDEP 6.0. Users Manual and Reference Guide. Version 6.0*. Econometric Software: New York, Belpport, 1992.
42. Coelli T. A guide to FRONTIER version 4.1: a computer program for stochastic frontier production and cost function estimation. *CEPA Working Paper*, Department of Econometrics, University of New England, Armidale NSW Australia, 1996.
43. Kodde DA, Palm FC. Wald criteria for jointly testing equality and inequality restrictions. *Econometrica* 1986; **54**: 1243–1248.
44. Linna M. Measuring hospital cost efficiency with panel data models. *Health Econ* 1998; **7**: 415–427.
45. Battese GE, Broca SS. Functional forms of stochastic frontier production functions, models for technical inefficiency effects: a comparative study for wheat farmers in Pakistan. *J Productivity Anal* 1997; **8**: 395–414.
46. Street A. How much confidence should we place in efficiency estimates? *Health Econ* 2003; **12**: 895–907.



The Cost Efficiency and Clinical Quality of Institutional Long-Term Care for the Elderly

JUHA LAINE*, MIIKA LINNA, ANJA NORO and UNTO HÄKKINEN

Centre for Health Economics at Stakes—CHESS, Helsinki, Finland

E-mail: juha.laine@stakes.fi

Received December 2003; accepted November 2004

Abstract. This study applied the stochastic frontier cost function with inefficiency effects to estimate the association between quality of care and cost efficiency in institutional long-term care wards for the elderly in Finland. We used several clinical quality indicators for indicating adverse care processes and outcomes, based on the Resident Assessment Instrument (RAI) / Minimum Data Set (MDS). Average cost inefficiency among the wards was 22%. We found an association between the clinical quality indicators and cost inefficiency. Higher prevalence of pressure ulcers was associated with higher costs, whereas the higher prevalence of use of depressants and hypnotics increased inefficiency.

Keywords: cost efficiency, stochastic frontier, long-term care, resident assessment instrument, minimum data set

1. Introduction

The association between costs and quality is a crucial issue in the production of health services, where it is usually assumed that low costs and high quality are conflicting objectives. Ensuring the quality of service may be costly, which can denote a cost–quality trade-off in the production and decision-making processes. According to Newhouse [1], quantity and quality are two commodities to which the decision-maker can allocate resources, and that a quantity–quality trade-off is present when resources are constrained. If there exists an association between costs or amount of output and quality, it definitely has implications relevant to policy and management. In that case we may ask what is the cost of quality or what are the implications of cost minimisation on the quality of care? Furthermore, productive efficiency may also explain the variation in costs between decision-making units. Hence, it is also relevant to consider factors that affect efficiency.

The effect of quality on (productive) efficiency in elderly care has been studied less than the cost–quality trade-off. However, it may well be that quality improvement can affect performance through altered activities and use of resources in the unit, which may decrease productive efficiency and cause higher costs. Producers can argue that processes for producing good quality care require certain types of staff mix and additional resources, which may prevent a unit from achieving technical and allocative efficiency, i.e. the best practice (efficient) production frontier. Producers may also argue that good outcomes of care are achieved only with better trained staff, i.e. at the cost of higher salary expenses. Analyses on the efficiency of elderly care in Finland—as well as in other countries—have

shown that nursing wards and facilities fail to achieve an efficient use of labor and capital resources for producing inpatient days [2]. However, the implication of the quality of care in the production process has remained somewhat unexplained.

In Finland, public discussion concerning problems in the quality of long-term care for the elderly has resembled the discussion in the US in the 1980s [3, 4]. Despite the broad debate, nationwide actions, such as government regulation aimed at ensuring and monitoring quality in the institutional care for the elderly, have been quite rare. Only the national recommendation for the adequate number of staff per patient and the recommendation to use quality monitoring systems are presented [4]. In addition, since well-defined and comparable quality measures are seldom available during decision-making processes, quality may in practice be subservient to price [5]. Public and private producers may have quality assurance programmes, quality principles or most often surveys on client satisfaction, which are used as indicators of quality. However, surveys capture only limited—and usually also incomparable—aspects of the quality of care. Service organisers (municipalities) in Finland are mainly forced to rely on very crude measures (e.g. costs per diem) as the basis for decisions. Due to the lack of comparable quality information, public producers can use their present quality level as an argument for their exceeded budgets and inefficiency.

Quality of care is a multidimensional concept, and there are several ways in which it can be operationalised and measured. According to Donabedian [6], quality measures can be regarded as structural (input), process, and outcome measures. In empirical studies, theoretical quality measures can be specified in various ways. For instance, if quality is measured by process measures, then quality can be regarded as a factor in the production process that may affect the efficiency of a unit. In this case, quality of care can be specified as an (in)efficiency

*Corresponding author.

effect in the transformation process of combining input(s) to yield output(s). Quality can also be defined as a final output or outcome, and it can be included as an independent variable in the cost function. In empirical efficiency studies in elderly care, quality has been specified variously, but theoretical discussion about the choice and implications for the quality measures used has usually been defective.

Our purpose was to consider the association between clinical quality of care and cost efficiency by applying an approach in which quality measures were included in the stochastic frontier cost function and as inefficiency effects. Quality of care was measured by clinical quality indicators, which describe adverse care processes and outcomes [7]. Case mix was adjusted by Resource Utilisation Groups (RUG-III) [8].

1.1. The Finnish elderly care system

In Finland, municipalities ($N = 446$) are responsible for providing the resident population with adequate social and health services. Financing is based mainly on state subsidies, local taxes and user fees. In practice, social and health services are usually produced by the municipalities, the voluntary or statutory coalition of municipalities or sometimes by private producers. Institutional elderly care is provided mainly in public residential homes and health-centre hospitals. Currently, the growing numbers of private service producers are concentrating more on the provision of care at home and of sheltered housing in the largest cities. Thus the comparison between profit and non-profit long-term care facilities is not yet a key issue in Finland, as it is in many other countries [9,10].

In general, residential homes provide basic care for residents whose physical or cognitive condition has deteriorated. The average length of stay in residential homes is about 3 years. Health-centre hospitals usually concentrate more on patients who need complex or medical care. The average length of stay in health-centre hospitals is about 2 years. The facility type can be associated with patient selection, unit size and other characteristics of a unit. Consequently, facility type is an important indicator of patient selection and of the operational environment of units in the Finnish elderly care system.

2. Literature review

Owing to the lack of quality measures, efficiency studies in particular are criticised for measurement bias [11]. Attempts have therefore been made to incorporate quality measures into theoretical studies [1,12,13] as well as into recent empirical efficiency studies on health services [13–16]. Most of the present literature on elderly care has dealt with efficiency, costs and the quality of care as separate issues and separate dimensions in performance. In addition, the focus has been mainly on the cost–quality trade-off—for instance, exploring how much quality costs—rather than the association between quality and efficiency. There are only a few studies on the economics of elderly care that have incorporated quality measures in the analysis, although these studies provide contradictory results.

Rosko et al. [9] argued that technical efficiency is not associated with the quality of care as measured by process and outcome indicators in US nursing homes. Efficiency was rather explained by managerial and environmental factors such as occupancy rate, ownership and payment source. Weech-Maldonado et al. [17] showed that nursing homes that produced better quality of care in terms of outcomes and processes reported better financial performance. However, nursing homes with better structural quality as measured by staffing ratios, experienced higher costs. Moreover, several studies have indicated a direct association between the use of labor inputs—for instance staffing ratios and staff mix—and quality of care [18–20].

Gertler and Waldman [13] found that the cost–quality trade-off depends on the economics of scale in quality in the nursing home industry, and a policy for quality improvements will be costly, though on the other hand a policy designed to reduce costs is possible with small reductions in quality. Blank and Eggink [14] applied Gertler's and Waldman's model in an analysis of Dutch nursing facilities. They concluded that quality was negatively related to input prices of personnel, implying potentially a loss of quality as input prices rise. Chou [10] addressed the issue that non-profit nursing homes produce better quality when asymmetric information is present. Knox et al. [21] found that quality influences costs marginally, and that efficiency differences reflect agency costs and differences in organisational goals. They suggested that as the quality of care increases, profitability increases too.

These results may indicate that distinct quality dimensions are in different ways associated with efficiency, costs and staffing. However, the empirical studies have been based on country-specific data, varying methods and model specifications, and different quality and case mix measures. It is hard to draw consistent conclusions from these studies. The results have indicated a potential association between costs or staffing and quality of care, but the association between quality and efficiency has remained inconclusive.

3. Data

The data for this study were drawn partly from our RAI (Resident Assessment instrument) database. Data have been collected in collaboration with participants in the project entitled “Benchmarking and Implementation of RAI in Elderly Care in Finland” [22]. Client-level assessments have been done using the RAI, which is a tool for improving the quality of nursing and care in all settings of elderly care. The RAI for nursing homes is based on the Minimum Data Set 2.0 (MDS 2.0) questionnaire, which is used to assess ward patients [3]. In Finland, each patient is assessed both at admission and then either every six months or when a distinct change in the patient's condition takes place. Registered nurses, assisted by other staff members and relatives, are usually responsible for the assessment. Several indices—such as that for dementia and for the case mix (Resource Utilisation Groups, RUG-III)—can be constructed from the MDS 2.0, thereby also providing

information that makes it possible to compose clinical quality indicators of adverse outcomes and adverse care processes [3,23–25].

Complete data pertaining to 113 wards in 38 facilities were collected from various sources. Information on total costs, salaries and full-time-equivalent person-years for each ward for the year 2001 was obtained from the finance offices of the facilities. Detailed instructions were given for cost collection and the allocation of overheads. A questionnaire filled in by the head nurses provided data on ward characteristics, such as staffing and occupancy rates. In this study we used case-mix index, gender distribution and quality indicators, which were aggregated for the ward level from individual-level data from the RAI database. There were about 3300 patients in the individual-level dataset.

We used the RUG-III classification for case-mix adjustment [8,26]. The case mix adjustment in this study was applied to the output variable, i.e. to the number of inpatient days. We calculated the patient days assigned to the RUG-III classification for each MDS assessment per episode of care in the ward, and calculated the total for each patient. In Finland, the 22-classification version of RUG-III is used [27]. The cost weight for different patients in 22-categories ranged from 0.48 to 2.52. The average cost weight is 1.0. The cost weights are derived from a study measuring the use of staff time.

About 30 clinical quality indicators were available from the RAI database [7,28]. In the first stage, these clinical indicators were considered from the conceptual viewpoint by experts in nursing to evaluate how they may be associated with the quality of care in residential homes and health-centre hospitals, and to ascertain whether they have been used in other studies as

quality measures. At the outset it was clear that some of these quality indicators were not appropriate for this study. This was the case if (i) the number of observations from which it was calculated were too low, (ii) there was not enough variation in the variable, (iii) the quality indicator was not applicable to both health-centre hospitals and residential homes, or (iv) the indicator was correlated strongly with other variables. In addition, the number of observations limited the number of (quality) variables used as independent variables. On the basis of reasoning, expert opinion, and the criteria listed above, four quality indicators indicating adverse care processes and outcomes were chosen for final analysis. The indicators were prevalence (%) of weekly use of depressants and hypnotics, prevalence of depression with no treatment, prevalence of pressure ulcers (risk adjusted) and prevalence of the use of restraints. The prevalence figures were the mean values of measurements made during three RAI assessment periods (cross-sections) falling between 16 September 2000 and 15 May 2002. They were calculated from several cross-sections in order to obtain more consistent quality measures and to diminish measurement errors and randomness.

A substantial number of ward characteristics—such as occupancy rate, the shift and team arrangements of the nursing staff, the number of purchased services, patient turnover, long-term or short-term ward, job rotation—were omitted from the final analysis, either because they were not systematically significant in the preliminary models or because they caused severe multicollinearity. Nevertheless, in this instance the focus of the study was on efficiency, and though these ward characteristics were omitted from the analysis, they may still be important in some other context. The descriptive statistics for the final variables are given in table 1.

Table 1
Data characteristics.

Variable name	Description	Mean (Std.)	LN mean (Std.)
Dependent			
Total costs in €	Total costs of the wards in € including capital investments and depreciations	1 114 591 (452 250)	13.85 (0.43)
Independent			
Case mix weighted inpatient days (output)	Calculation of the patient days associated with the RUG-III classification for each MDS assessment within an episode of care (=the ward's total case mix weighted patient days).	10 400 (4 533)	9.19 (0.39)
Average wage rate	Average annual salary of the staff in €	36 078 (6 215)	8.69 (0.19)
Facility type	0 = Residential home, 1 = Health-centre hospital	–	–
Prevalence of depression without treatment ¹	%, outcome/process measure	14 (7)	–
Prevalence of pressure ulcers, high risk ¹	%, high risk adjusted outcome measure	11 (66)	–
Prevalence of the weekly use of depressants and hypnotics ¹	%, process measure	60 (14)	–
Prevalence of restraints ¹	%, process measure	18 (15)	–

¹Higher prevalence (%) indicates poorer quality.

4. Methods and model specification

4.1. Method

There are two main approaches to the parametric analysis of inefficiency effects: the traditional two-stage approach and the technical inefficiency effect estimation (single-stage approach). Kumbhakar, Ghosh and McGukin [29], and Reifschneider and Stevenson [30], proposed stochastic frontier models in which the inefficiency effects are expressed as an explicit function of a vector of firm-specific variables and a random error. They argued that the parameter estimates in the second-stage regression of the two-stage approach are not consistent, and that this can be avoided by using single-stage estimation. These arguments have met with wide acceptance, and have been further supported by Wang and Schmidt [31] and Battese and Coelli [32], who also concluded that the two-stage approach can lead to serious bias. A model for cost inefficiency effects in the stochastic frontier cost function for cross-sectional data can be defined as

$$C_i = C(y, w) + v_i + u_i, \quad (1)$$

where

- C_i = total costs,
- y = outputs
- w = input prices
- v_i = a random error distributed as $\sim N[0, \sigma_v^2]$ and
- u_i = a inefficiency term.

The inefficiency component u_i is comprised of non-negative random variables (inefficiency effects) that are assumed to be independently distributed as truncations of the $N(\delta Z_i, \sigma_u^2)$ distribution. Inefficiency effects, denoted Z_i , arise by truncation at zero of the normal distribution with mean μ_i ($\mu_i = \delta Z$) and variance σ_u^2 . The δ -coefficients are unknown parameters to be estimated together with variance parameters $\sigma^2 = \sigma_v^2 + \sigma_u^2$ and $\gamma = \sigma_u^2 / (\sigma_v^2 + \sigma_u^2)$.

Another important question relates to the choice of other independent variables and inefficiency effect variables, where there seems to be several practices related to the incorporation of quality measures in the analysis. In empirical studies, quality measures have been entered into the (stochastic frontier) cost function [21,33–35] or as separate inefficiency effects [15,36–38] or in both the frontier function and as inefficiency effects [39,40]. In some studies quality measures are also considerably omitted [41]. The variables included in the cost function affect the best practice frontier and costs, while the variables in the inefficiency effect component affect the distance from the frontier. These specifications are based on different theoretical foundations, and thus lead to different interpretations. In this study output and outcome measures, i.e. two outcome quality measures and inpatient days, were included in the stochastic cost frontier. The two process quality measures, which may cause inefficiency and good or poor outcomes, were specified as inefficiency effects.¹

¹ This model specification was suggested by the reviewer.

4.2. Model specification

The specification of a functional form for the cost function is essential in the empirical implementation of the cost frontier approach. In the first stage we applied a translog specification of the cost function, which is a widely used specification due to its flexibility. The second order and cross-terms of the output and input price variables were not significant. Thus, we applied Box-Cox transformation, which is $y^{(\lambda)} = (y^\lambda - 1)/\lambda$, to the output variable to allow flexibility in the output variable. Box-Cox analysis was performed using LIMDEP software [42]. The Box-Cox analysis suggested logarithmic transformation of the output variable. The logarithmic specification of the model also increased the degrees of freedom compared to the translog specification. The standard assumption of linear homogeneity in input prices was not imposed, as there were no satisfactory means to measure the true prices of equipment and materials (capital). in the public units. Hence, the input price variable included only labour and it was named as the average wage rate. Thus, our average wage rate variable is a somewhat incomplete proxy of input prices. The average wage rate was at an aggregated level, so it was not possible to distinguish the cost of labour for each of the different staff categories.

In addition, the dummy for the type of facility was included in the cost frontier function as indicators of patient selection and operational environment in different type of facilities. It is possible that facility type is associated with the best practice frontier and different production technology, although the case-mix adjustment was applied to the output variable. Two quality indicators indicating poor outcomes of care were also included in the cost frontier function. The final functional form of the stochastic cost frontier was specified as:

$$\begin{aligned} \ln(TC_i) = & \beta_0 + \beta_1 \ln(CMOUTPUT_i) + \beta_2 \ln(WAGE_i) \\ & + \beta_3(FACILITY_i) + \beta_4(DEPRESS_i) \\ & + \beta_5(PRESS_i) + v_i + u_i \end{aligned} \quad (2)$$

where TC denotes the logarithm of the total costs, CMOUTPUT is the logarithm of inpatient days weighted by case mix, WAGE is the logarithm of the average wage rate, FACILITY is a dummy for the type of facility (residential home or health centre hospital), DEPRESS is the prevalence of depression without treatment, PRESS is the prevalence of pressure ulcers (risk adjusted), v_i is statistical noise and u_i is the inefficiency component. The quality of care was the subject of this study, so two process quality indicators were included as inefficiency effects. The final model for the cost inefficiency effect equation was defined by:

$$\mu_i = \delta_0 + \delta_1(MED_i) + \delta_2(REST_i) + w_i, \quad (3)$$

where MED is the weekly use of depressants and hypnotics, REST is the prevalence of the use of restraints and w_i is defined by the truncation of the normal distribution with zero mean and variance, σ^2 , such that μ_i is non-negative.

Table 2
Generalised likelihood-ratio tests of null hypotheses for parameters of the stochastic frontier cost function.¹

Null hypothesis	χ^2 -statistic	$\chi^2_{0.95}$ -value	Decision
Stochastic frontier			
$H_0: \gamma = \delta_0 = \delta_1 = \delta_2 = 0$	9.88	8.76	Reject H_0
$H_0: \delta_1 = \delta_2 = 0$	6.27	4.30	Reject H_0
$H_0: \delta_0 = 0$	0.56	12.71	Accept H_0

¹ The critical values for the hypothesis $\gamma = 0$ were taken from table 1 in Kodde and Palm [44].

The stochastic frontier cost function and inefficiency effects were estimated using FRONTIER 4.1 software written by T. Coelli [43]. The program uses a three-step procedure for estimating the maximum likelihood estimates of the parameters of a stochastic frontier cost function.

5. Results

The mean cost inefficiency was 22%. The cost inefficiency was less than 10% in 23% of the wards, while in 27% of the wards, cost inefficiency was at least 30%. Several restrictions of the specified inefficiency effect model were tested by applying a likelihood ratio test (table 2). Two of the following null hypotheses were rejected at the 0.05 level: (i) inefficiency effect variables do not jointly explain inefficiency differences and (ii) quality as an inefficiency effect is absent from the model. The hypothesis that inefficiency effects do not have an intercept parameter was, however, accepted. Hence, the preferred model was estimated without δ_0 . The decomposed error term was specified to follow half normal distribution. The maximum-likelihood estimates of the parameters in the stochastic frontier cost functions and the estimates of inefficiency effects are given in table 3.

The estimated coefficients of case mix weighted output, average wage rate, and facility type were all positive and significant. Health-centre hospitals seemed to have higher costs than residential homes, which may be an outcome of the higher number of professional staff members, higher labour and allocated capital costs incurred in health-centre hospital wards and a different operational environment. A higher prevalence of poor outcome quality in terms of the prevalence of pressure ulcers (risk-adjusted) was positively and significantly associated with higher costs. Calculated elasticities for both of the outcome quality indicators were 0.02. This result indicates that a 10% increase in poor outcomes increases total costs by 2%. A positive value of the delta parameter (inefficiency effect) in table 3 implies a positive relationship with inefficiency. Poor process quality in terms of the weekly use of depressants and hypnotics was positively associated with inefficiency, whereas the prevalence of the use of restraints was not significantly associated with inefficiency.

The magnitude of the effect of changes in process quality measures upon cost efficiency was addressed by calculating marginal effects associated with inefficiency effects. The cal-

Table 3
Maximum likelihood estimates for parameters of the stochastic frontier cost function with inefficiency effects models involving ward-specific variables.¹

Variable	Parameter	Model estimates
Stochastic frontier		
Constant	β_0	0.61 (0.94)
Logarithm of inpatient days weighted by case mix	β_1	0.91 (32.89)
Logarithm of average wage rate	β_2	0.52 (9.34)
Facility type	β_3	0.30 (12.64)
Prevalence of depression with no treatment	β_4	0.0015 (1.09)
Prevalence of 1.-4. stage pressure ulcers	β_5	0.0019 (2.05)
Inefficiency effects		
Constant	δ_0	—
Prevalence of use of depressants and hypnotics	δ_1	0.0027 (3.28)
Prevalence of use of restraints	δ_2	−0.0001 (−0.05)
Variance Parameters		
	σ^2_ε	0.021 (4.03)
	γ	0.999 (1642.41)
Likelihood function		94.74

¹ Asymptotic *t*-test in parenthesis.

Table 4
Marginal effects of the inefficiency effects.

Variable	Parameter	Marginal effect
Prevalence of use of depressants and hypnotics	δ_1	0.0011
Prevalence of use of restraints	δ_2	−0.00004

culation was performed by following a formula derived by Frame and Coelli [45]. Marginal effects of the process quality measures and occupancy rate proved to be insubstantial (table 4).

In order to further elucidate more the relationship between cost inefficiency and inefficiency effects, the change in predicted cost efficiency when a ward's process quality changes from the first decile or quartile to the tenth decile or fourth quartile of the process quality distributions is illustrated in figure 1. Calculations were performed separately for each of the tested variables, keeping the other variables constant using the parameter estimates of the final model. The variation in quality distributions was considerable. For instance, the prevalence (%) of weekly use of depressants and hypnotics was 35 in the lowest decile, on average, but about 83 in the highest decile. Despite this, the predicted change in inefficiency seemed to be about 0.13 efficiency units when a ward moved from the lowest to the highest decile in the quality distribution. The mean efficiency score was 1.22.

For the sensitivity analysis, instead of using a high-risk-adjusted indicator of the prevalence of pressure ulcers in the cost function, we also tested non-risk-adjusted and low-risk-adjusted indicators. The mean inefficiency as well as the parameter estimates for the output, average wage and facility type remained unchanged. However, the statistical significance of

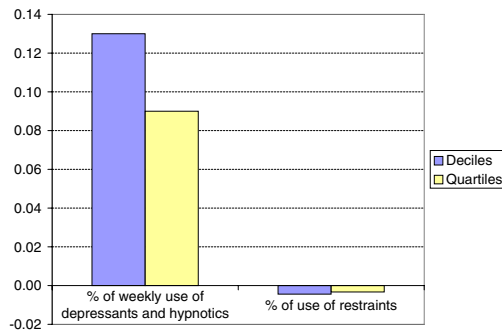


Figure 1. The effect on inefficiency of transition from the lowest to the highest deciles and quartiles in the distribution of quality indicators.

the other outcome and process quality indicators varied moderately. The sign of the quality variables were otherwise consistent. Despite the fact that the preferred model was somewhat sensitive to the set of the quality indicators used, all tested model specifications led quite systematically to the same inference.

6. Discussion and conclusion

There was marked variation in clinical quality of care and cost efficiency in the elderly care wards, which indicated clear improvement potential in the overall performance of the units. The mean cost inefficiency of all wards was 22%, which was somewhat similar to those found in earlier studies on Finnish long-term care [46]. Poor outcome quality in terms of the prevalence of pressure ulcers was associated with higher costs. Moreover, higher weekly use of depressants and hypnotics as a measure of poor process quality was associated with inefficiency. The results of this study suggest that poor quality of care may be associated more clearly with higher costs than cost inefficiency, which is also a result consistent with the previous studies [9,14,17]. On the whole, the quality of care had a relatively weak effects on costs and inefficiency compared to the other factors used in the cost function, but the results indicate clearly that quality should not be omitted from efficiency measurement.

There is no composite quality measure for overall quality in institutional elderly care [47]. Indicators from all quality dimensions—such as subjective measures of quality experienced by patients or objective measures of structure, process or outcomes—are seldom available in empirical studies. There is therefore always the risk of bias stemming from the omission of quality dimensions. More effort is therefore needed to take multidimensional quality into account in performance measurement. Our study gave some new insight into how the association between cost efficiency and multidimensional quality can be explored.

In institutional elderly care there are many situations in which the use of clinical quality indicators may be more feasible than subjective quality measures, which are, however, generally used. Long-term care patients' cognition is usually impaired, and relatives may be absent, and thus patients cannot necessarily evaluate quality or express a reliable opinion. On the other hand, objective quality measures also have deficiencies. They may suffer from selection bias and ascertainment bias, depending on the method of measurement [7,48]. Non-randomised selection of patients from the wards and facilities may have an effect on client structure, and this can have an effect on clinical quality indicators. Hence, some clinical quality measures, especially some non-risk-adjusted measures, may indicate case-mix instead of quality. One weakness of this study was that only cross-sectional data were available, though the mean of the quality indicators was calculated from three cross-sections over an 18-month period in order to diminish randomness and measurement errors. Efficiency estimates would also be more consistent if longitudinal data were used [49].

Some limitations also need to be taken into consideration when quality measurement is incorporated into the parametric analysis of efficiency. Although there were almost 30 clinical indicators available from the RAI database, we could not include all of them in the parametric analysis because of the quite low number of observations and the correlation between variables. Therefore, the process and outcome quality indicators used in this study were chosen on a conceptual and statistical basis. Changes in quality or in ward characteristics may also be associated with efficiency in a more complex way than our analysis showed. For instance, changes in the prevalence of restraints can have an effect on the prevalence of outcome quality indicators, staffing or other factors in the operational environment, at least in the long-run. Moreover, changes in quality level may well have an influence on the effectiveness of care, on client satisfaction, and on the reputation of a ward. These aspects were not considered in this study.

We employed two types of quality measures—process and outcome—in our inefficiency effect model specification. The results indicates that although quality is an important factor in efficiency and performance measurement, the model specification needs to be considered carefully when quality measures are incorporated into parametric efficiency analysis. More work is needed to develop an appropriate modelling framework to take into account several quality aspects and to understand the mechanism for how outcome quality affects costs and how process quality affects inefficiency in elderly care.

Acknowledgements

This study has been funded by the Academy of Finland (Academy decision notification number 70034), the National Research and Development Centre for Welfare and Health (STAKES) and the Doctoral Programme in Social and Health Services Management and Economics of the University of Kuopio. We are grateful to Professor Tim Coelli for the Frontier

4.1c software. Grateful acknowledgement is also given to interRAI.

References

- [1] J. Newhouse, Toward a theory of nonprofit institutions: An economic model of a hospitals, *American Economic Review* 60 (1970) 64–74.
- [2] J. Laine, Pitkääikaishoidon tehokkuus ja laatu vuosina 2000–2002, in R. Hjerpe, A. Kangasharju and R. Vuorento (eds.), *Kunnalliset palvelut—Terveyden- ja vanhustenhuollon tuottavuus*. Helsinki: VATT, 2003.
- [3] J.N. Morris, C. Hawes, B.E. Fries, C.D. Phillips, V. Mor, S. Katz, K. Murphy, M.L. Drugovich and A.S. Friedlob, Designing the national resident assessment instrument for nursing homes, *The Gerontologist* 30 (1990) 293–307.
- [4] M. Vaarama, J. Luomahaara, A. Peiponen and P. Voutilainen, Koko kunta ikääntyneiden asialle. Näkökulmia ikääntyneiden itsenäisen selviytymisen sekä hoidon ja palvelun kehittämiseen, in Stakes, (ed), *Raportteja*. Helsinki: Stakes, 2001.
- [5] J. Laine, Selvitys sosiaalihuollon erityispalveluiden tuotantomalleista ja sopimusmenettelyistä. Helsinki: Stakes, 1999.
- [6] A. Donabedian, The quality of care? How can it be assessed? *JAMA* 260 (1988) 1743–1748.
- [7] D.R. Zimmermann, S.L. Karon and G.E.A. Arling, Development and testing of nursing home quality indicators, *Health Care Finance Review* 16 (1995) 107–140.
- [8] B.E. Fries, J.R. Cooney and M. Leo, Resource utilization groups. A patient classification system for long-term care, *Medical Care* 23 (1985) 110–122.
- [9] M. Rosko, J. Chilingirian, J. Zinn and W. Aaronson, The effects of ownership, operating environment, and strategic choices on nursing home efficiency, *Medical Care* 33 (1995) 1001–1021.
- [10] S.-Y. Chou, Asymmetric information, ownership and quality of care: An empirical analysis of nursing homes, *Journal of Health Economics* 21 (2002) 293–311.
- [11] J. Newhouse, Frontier estimation: How useful a tool for health economics?, *Journal of Health Economics* 13 (1994) 317–322.
- [12] A.O. Hirschman, *Exit, Voice, and Loyalty*. Cambridge, Massachusetts: Harvard University Press, 1970.
- [13] P.J. Gertler and D.M. Waldman, Quality-adjusted cost functions and policy evaluation in the nursing home industry, *Journal of Political Economy* 100 (1992) 1232–1256.
- [14] J.L.T. Blank and E. Eggink, A quality-adjusted cost function in a regulated industry: The case of dutch nursing homes, *Health Care Management Science* 4 (2001) 201–211.
- [15] M. Rosko, Cost efficiency of us hospitals: A stochastic frontier approach, *Health Economics* 10 (2001) 539–551.
- [16] O. Olesen and N. Petersen, Incorporating quality into data envelopment analysis: A stochastic dominance approach, *International Journal of Production Economics* 39 (1995) 117–35.
- [17] R.J. Weech-Maldonado, G. Neff and V. Mor, Does quality of care lead to better financial performance? The case of the nursing home industry., *Health Care Management Review* 28 (2003) 201–216.
- [18] M.E. Dellefield, The relationship between nurse staffing in nursing homes and quality indicators, *Journal of Gerontological Nursing* 27 (2000) 14–28.
- [19] C. Harrington, D. Zimmermann, S.L. Karon, J. Robinson and P. Beutel, Nursing home staffing and its relationship to deficiencies, *Journal of Gerontology: Social Sciences* 55 (2000) 278–287.
- [20] GAO, *Nursing Homes—Quality of Care More Related to Staffing than Spending*, in GAO, T. G. A. O.-. ed., *GAO Report*. Washington: The General Accounting Office—GAO, 2002.
- [21] K.J. Knox, E.C. Blankmeyer and J.R. Stutzman, Organizational efficiency and quality in texas nursing facilities, *Health Care Management Science* 6 (2003) 175–188.
- [22] A. Noro, H. Finne-Soveri, M. Björkgren, U. Häkkinen, J. Laine, P. Vähäkangas, S. Kerppilä and R. Storbacka, *RAI -tietojärjestelmän esittely ja kokemuksia pitkäaikaishoidon benchmarkingista*. RAI -raportti 1/2000. Helsinki: Stakes, 2001.
- [23] C. Hawes, J. Morris, C. Phillips, V. Mor, B. Fries and S. Nonemaker, Reliability estimates for the minimum data set for nursing home resident assessment and care screening, *The Gerontologist* 35 (1995) 172–178.
- [24] C. Hawes, J. Morris, C. Phillips, B. Fries, C. Murphy and V. Mor, Development of the nursing home resident assessment instrument in the USA, *Age and Ageing* 26–S2 (1997) 19–25.
- [25] C. Murphy and J. Morris, The resident assessment instrument: Implications for quality, reimbursement and research., *Generations* 19 (1995) 43–46.
- [26] B.E. Fries, D.P. Schneider, W.J. Foley, M. Gavazzi, R. Burke, and E. Cornelius, Refining a case-mix measure for nursing homes: Resource utilization groups (RUG-III), *Medical Care* 32 (1995) 668–685.
- [27] M. Björkgren, U. Häkkinen, H. Finne-Soveri and B. Fries, Validity and reliability of Resource Utilization Groups (RUG-III) in Finnish long-term care facilities, *Scandinavian Journal of Public Health* (1999) 228–234.
- [28] S. Karon and D. Zimmermann, Using indicators to structure quality improvement initiatives in long-term care, *Quality Management in Health Care* 4 (1996) 54–66.
- [29] S. Kumbhakar, S. Ghosh and J. McGuckin, A generalized production frontier approach for estimating determinants of inefficiency in U.S. dairy farms., *Journal of Business & Economic Statistics* 9 (1991) 279–286.
- [30] D. Reifschneider and R. Stevenson, Systematic departures from the frontier: A framework for the analysis of firm inefficiency, *International Economic Review* 32 (1991) 715–723.
- [31] H.-J. Wang and P. Schmidt, One-step and two-step estimation of the effects of exogenous variables on technical efficiency levels, *Journal of Productivity Analysis* 18 (2002) 129–144.
- [32] G.E. Battese and T.J. Coelli, A stochastic frontier production function incorporating a model for technical inefficiency effects, *Working Papers in Econometrics and Applied Statistics* (1993).
- [33] L. Crivelli, M. Filippini and D. Lunati, Regulation, ownership and efficiency in the swiss nursing home industry, *International Journal of Health Care Finance and Economics* 2 (2002) 79–97.
- [34] S.T. Folland and R.A. Hofer, How reliable are hospital efficiency estimates? Exploiting the dual to homothetic production, *Health Economics* 10 (2001) 683–698.
- [35] S. Zuckerman, J. Hadley and L. Iezzoni, Measuring hospital efficiency with frontier cost functions, *Journal of Health Economics* 13 (1994) 255–280.
- [36] T. Li and R. Rosenman, Cost inefficiency in washington hospitals: A stochastic frontier approach using panel data, *Health Care Management Science* 4 (2001) 73–81.
- [37] M. Linna, Measuring hospital cost efficiency with panel data models, *Health Economics* 7 (1998) 415–427.
- [38] G.E. Battese and T.J. Coelli, A model for technical inefficiency effects in stochastic frontier production function for panel data, *Empirical Economics* 20 (1995) 325–332.
- [39] C.J. Huang and J.T. Liu, A non-neutral stochastic frontier production function, *Journal of Productivity Analysis* 5 (1994) 171–180.
- [40] S. Ngwenya, G. Battese and E. Fleming, The relationship between farm size and the technical inefficiency of production of wheat farmers in the eastern free state province of South Africa, *Agrekon* 36 (1997) 283–301.
- [41] M.D. Rosko, Performance of U.S. Teaching Hospitals: A panel analysis of cost inefficiency, *Health Care Management Science* 7 (2004) 7–16.
- [42] W. Greene, *LIMDEP 6.0. User's Manual and Reference Guide*. Version 6.0. New York: Belpor: Econometric Software, Inc., 1992.
- [43] T. Coelli, *A Guide to FRONTIER version 4.1: A computer program for stochastic frontier production and cost function estimation*, CEPA Working Paper: Department of Econometrics, University of New England, Armidale NSW Australia, 1996.

- [44] D.A. Kodde and F.C. Palm, Wald criteria for jointly testing equality and inequality restrictions, *Econometrica* 54 (1986) 1243–1248.
- [45] S.W. Frame and T.J. Coelli, U.S. Financial services consolidation: The case of corporate credit unions, *Review of Industrial Organization* 18 (2001) 229–242.
- [46] M. Björkgren, U. Häkkinen and M. Linna, Measuring efficiency of long-term care units in finland, *Health Care Management Science* 4 (2001) 193–200.
- [47] N. Ikegami, J.P. Hirdes and I. Carpenter, Measuring the quality of long-term care in institutional and community setting, in OECD (ed.), *Measuring UP: Improving Health System Performance in OECD Countries*. Paris: OECD, 2002.
- [48] CMS. Identification and Evaluation of Existing Quality Indicators that are Appropriate for Use in Long-Term Care Setting, in Inc, A. A. (ed.), *The Center for Medicare and Medicaid Services (CMS)*, 2001.
- [49] A. Street, How much confidence should we place in efficiency estimates?, *Health Economics* 12 (2003) 895–907.

Patient- and ward-level determinants of nursing time in nursing facilities

Juha Laine, Anja Noro, Harriet Finne-Soveri & Unto Häkkinen
Centre for Health Economics at Stakes – CHESS, Helsinki, Finland

Abstract

Objectives: To explore the patient- and ward-level determinants of wage-adjusted nursing time in long-term care wards for the elderly with a view to improving efficiency of the use of labour resources.

Methods: The wage-adjusted nursing time given to patients was obtained from a time measurement study. Patient characteristics were based on the Minimum Data Set 2.0 of the Resident Assessment Instrument for nursing homes. Ordinary least-squares regression analysis and multilevel modelling were used to disentangle the effect of patient- and ward-level factors on nursing time.

Results: A significant difference in wage-adjusted nursing time between wards was detected, which was partly explained by characteristics of patient in wards. The combination of patients' physical functioning and cognition and the RUG-III/22 resource use classification explained 20–25% of patients' nursing time over a 24-h period. Variables related to operational environment of the ward did not explain differences in wage-adjusted nursing time once the patient profile of the ward had been controlled for. The results also showed notable unmeasured patient and ward level effects, inefficiency and randomness in the allocation of nursing time.

Conclusions: By improving the allocation and use of labour resources, the substantial variation in nursing time differences between wards could be diminished. Managers should allocate their staff primarily according to patients' resource needs.

Introduction

Staff and the care time they provide are the most important resources in institutional long-term care for the elderly. Labour costs account for approximately 70–80 per cent of total costs in Finland. Nursing staffing and time are directly associated with the quality of care in nursing homes and hospitals.^{1–4} This highlights the importance of exploring

the determinants of nursing time. Residents in nursing facilities tend to be increasingly frail,⁵ which emphasizes the high degree of professional skills required by trained workers. Unfortunately, the mean age of nursing personnel is rising,⁶ salaries are not competitive,⁷ and an aging population increasingly demand more help at home and in long-term care institutions.⁶ Hence, trained personnel are becoming a scarce resource in many

Correspondence: Juha Laine, STAKES, Lintulahdenkuja 4, PO Box 220 FIN-00531 Helsinki, Finland.

Tel: +358 (0)9 39672303, Fax: +358 (0)9 39672485, E-mail: juha.laine@stakes.fi

long-term facilities. Finland already has quite a low number of staff per resident in long-term care facilities compared with other Nordic countries.⁸

Consequently, ensuring an adequate number of trained staff in long-term care facilities is a challenge for the future. However, it is not only the number of staff that matters. More staff in total do not necessarily guarantee better care, nor fair or effective allocation of labour resources. Both ensuring adequate staffing and effective utilization of staff are key issues in care of the elderly.

Several case-mix classifications are based on an assessment of the patient characteristics, of which physical functioning seems to be the most important element.^{9–11} Arling and Williams¹² showed that cognition and especially physical functioning have a substantial effect on resident-specific care time in nursing homes. Long-term care patients are, however, clustered within a ward and a facility, in which case the nursing practices, operational management, premises, and efficiency may also affect the nursing time provided and the allocation of labour resources.

Analyses of the efficiency of elderly care in Finland have shown that wards fail to achieve an effective use of labour and capital resources and that there is variation in the staff-mix and employee per resident ratio between units.¹³ Ward dummy variables were included by Björkgren et al.¹⁴ in the validation study of RUG-III (Resource Utilization Groups) for long-term care patients in Finland in order to indicate the efficiency differences between wards. The analysis showed that ward dummy variables were significant and that they improved the patient level case-mix model. Takahashi et al.¹⁵ explored the relationship between case-mix and the resource allocation in long-term care facilities in Japan. They

concluded that facility managers decide the staff allocation not only with reference to the patient mix of units, but also the management policy and the physical characteristics of a unit. However, once the staff are assigned to a unit, they tend to allocate their time mainly on the basis of the case-mix of patients.

The effects of ward and facility level factors are not usually reported or explicitly measured when the determinants of patient's nursing time are explored, even though a broader approach may be warranted. If the operational environment is associated with nursing time, the management should be aware of potential effects on outcomes of care and on the allocation of staff. For instance, in larger units or in units where the proportion of single rooms is low, patients may be given less nursing time compared to more home-like units. Alternatively, the nursing time given to patients may increase as a result of the shift and team arrangements of staff, a lower level of medications and a ward specialization in the care of particular types of patients. Most of these environmental characteristics can be manipulated, so they should be essential factors in the decision-making process. The potential effect of external environmental characteristics may indicate that staff time is not or cannot merely be allocated according to patient need and the patient profile of the unit.

The purpose of this study is to analyse the determinants of wage-adjusted nursing time in health centre hospitals and residential homes. Particular attention is paid to ward-level factors, which describe the different operational environments of wards. We explore how these factors affect wage-adjusted nursing time after controlling for the patient profile of a ward.

Design and methods

Data

Data for this study were collected as part of the project entitled "Benchmarking and Implementation of RAI in Elderly Care in Finland".¹⁶ Data were also included from the time measurement study involving 86 wards in 19 facilities during November 2002, which made use of the conventional time study method used in the RUG-III validation studies.^{14,17} The nursing staff ($n = 1031$) (i.e. registered nurses, licensed practical nurses and nursing aides) used time sheets to record the direct time spent with a resident and time spent in general tasks over a 24-hour period. Direct care time was defined as the time spent in hands on care and all time spent on care planning, when that time could be allocated to a specific resident. General or ward time was defined as time and tasks that could not be allocated to a specific resident. The direct time of therapists and other auxiliary staff ($n \approx 300$) was recorded over a seven-day period. Informal care time provided by family members and friends ($n \approx 1170$) was included only to the extent that it replaced formal care time. The staff estimated from the informal carers' time sheets the amount of time the carers gave to tasks, which would otherwise be carried out by the staff, such as toileting, over a seven-day period. Patients' total nursing time over a 24-h period was defined as the sum of these care times weighted by cost using the wage rates of the respective staff categories.⁷ For instance, the weight for registered nurses was 1.00 and 0.85 for licensed practical nurses. Since there is not an objective salary weight for informal carers, the informal care time was given a weighting of 0.5 (while a weight 0.75 was also tested). Hence, the nursing time variable also represents a cost measure.¹¹

The patient characteristics were derived from the Minimum Data Set 2.0 (MDS) of the Resident Assessment Instrument for nursing homes.¹⁸ We used RUG-III/22 resource use groups for characterizing different types of patients.¹⁴ In order to explore in more detail the effect of patient-level factors, several MDS items were also tested separately in the analysis in order to reveal any non-linear or reverse association between different patient characteristics and nursing time that might be unobserved if only the RUG-III/22 groups were used. Separate scales and variables were derived from the MDS and chosen on the basis of clinical reasoning: activities in daily living (hierarchical ADL, scale 0–6) to represent physical functions,¹⁹ cognitive condition (CPS, scale 0–6),²⁰ daily pain,²¹ depression symptoms (DRS, scale 0–14),²² unbalanced relationship with others, age, nursing rehabilitation and admission date. We also considered the stability of health, physical aggressiveness and gender, but they were not associated with nursing time, so we excluded them from the final analysis. Staff time measurements and MDS assessments were combined if patients or their relatives had given written consent.

Ward characteristics describing the operational environment of the wards were measured using a questionnaire targeted at head nurses. The ward characteristics used in the final analysis were: nursing hours as a proportion of the staff's total working hours, the unit size, the average number of medications during the previous seven days, the proportion of rooms with *en suite* bathroom, the proportion of single rooms, the shift and team arrangements of the nursing staff, the ward's specialisation in psycho-geriatric patient care, and the facility type. The descriptive statistics of

the variables used in the final analysis are given in Table 1. Time measurement and MDS assessment data were combined

with the ward-level data. The final study sample included 1787 residents, 68 wards, and 16 facilities.

TABLE 1. Descriptive statistics of the data

Variable Name	Description	Mean (Std.)	Minimum – maximum
Dependent variable			
Nursing time in minutes over a 24-hour period	The total of nursing time in minutes per resident per day weighted by the salary of different staff categories. Nursing time is the direct or indirect care time that could be allocated to a specific resident.	99 (43)	2 - 424
Patient-level variables			
RUG-III/22 classification	Dummy variables representing 22 different RUG-III groups	-	0-1
Activities of Daily Living (ADL) scale and Cognitive performance (CPS) scale.	The integrated variable of the original ADL and CPS scales (ranged 0-6): 0 = Totally independent (ADL 0-2) / intact (CPS 0-2) 1 = Moderately dependent (ADL 3-4) / intact (CPS 0-2) 2 = Fully dependent (ADL 5-6) / intact (CPS 0-2) 3 = Totally independent (ADL 0-2) / moderately impaired (CPS 3-4) 4 = Moderately dependent (ADL 3-4) / moderately impaired (CPS 3-4) 5 = Fully dependent (ADL 5-6) / moderately impaired (CPS 3-4) 6 = Totally independent (ADL 0-2) / impaired (CPS 5-6) 7 = Moderately dependent (ADL 3-4) / impaired (CPS 5-6) 8 = Fully dependent (ADL 5-6) / impaired (CPS 5-6)	0.09 0.09 0.07 0.05 0.11 0.13 0.02 0.09 0.35	0-1 0-1 0-1 0-1 0-1 0-1 0-1 0-1 0-1
Unbalanced relationship with others	Composed from the MDS 2.0 questionnaire's part F2 (psycho-social well-being). 0 = No problems with relationships with others 1 = Problems with relationships with others (such as hostility or conflict with others)	0.37	0-1
Daily pain	0 = No, 1 = Yes	0.21	0-1
Depression symptoms (DRS)	0 = No, 1 = Yes (DRS scale score >5)	0.12	0-1
Nursing rehabilitation	0 = No nursing rehabilitation within 7 days 1 = Nursing rehabilitation within 7 days	0.77	0-1
Admission	Days since admission to the ward	1030 (1098)	1 - 9020
Age	0 = < 65 1 = 65 - 74 2 = 75 - 84 3 = 85 -	80 (10)	22 - 104
Ward-level variables			
The proportion of nursing hours	% of the total work hours. Nursing hours is the direct or indirect time that could be allocated to a specific resident.	49 (7)	31 - 62
Unit size	The number of the beds	30 (8)	8- 44
Medications	Number of medications during past 7 days	7.24 (1.29)	3.00 - 10.08
The proportion of rooms with own bathroom	%	63 (40)	0 - 100
The proportion of single rooms	%	41 (28)	0 - 100
Shift and team-arrangements of the nursing staff	0 = No shift and team arrangements 1 = Shift and team arrangements	0.78	0-1
Ward's specialisation	0 = Mixed patients 1 = Psycho-geriatric patients (at least 60% of patients belonging to RUG-III/22 groups IA, IB, PC, PD, BA, BB)	0.45	0-1
Facility type	0 = Residential home 1 = Health centre hospital	0.52	0-1

Methods

The analysis was conducted in two stages. In the first stage we used ordinary least-squares regression (OLS) in order to describe the ward effect on average wage-adjusted nursing time in the wards. The *observed* average wage-adjusted nursing time per patient was compared with the RUG-III/22 *predicted* average wage-adjusted nursing time per patient in the wards. In the second stage, patients' nursing time was explained by the patient characteristics and ward dummies using OLS. Furthermore, we also applied multilevel modelling, which is a widely used technique to analyze data that fall naturally into hierarchical structures.^{23,24,25} Multilevel modelling allowed an exploration of the extent to which effects of ward characteristics determined nursing time after controlling patient characteristics. In the case of hierarchical structures, an OLS may lead to an underestimation of standard errors and thereby lead to misleading inferences.

The basic multilevel model for two-level hierarchical data is

$$y_{ij} = \alpha + \beta_p x_{ij} + (u_j + e_{ij})$$

The difference with OLS is that the error term is divided, so that u_j is a higher level residual and e_{ij} is a lower level residual. Subscript j refers, for instance to the ward, and subscript i to the patient. The disturbance terms are random variables distributed normally with mean 0 and variance σ^2 . The model can be extended to several hierarchies.

In this study, we used a variance component (multilevel) model, in which explanatory variables were measured at patient and ward levels (Table 1). The

models were estimated using MLwiN 1.1 software.²⁶ The multilevel models are sensitive to the standard assumptions of OLS such as multicollinearity and model specification. Severe multicollinearity was not detected in the data. With regard to the continuous variables, the non-linear relationships were tested by adding squared variables to the model. No significant interactions were identified.

The ADL scale and the CPS were integrated in order to compose a clinically reasonable and novel measure of physical and cognitive performance, which was divided into three categories: (i) good (0-2), (ii) moderate (3-4) and (iii) poor (5-6). The classification is described in Table 1.

Results

There was a marked variation in wage-adjusted nursing time between patients and wards. At the patient level the minimum was two and the maximum 424 minutes per patient per 24 h. At the ward level, the average wage-adjusted nursing time varied from 66 to 137 minutes per day. The mean was 99 minutes.

In the first stage, we estimated the OLS model with RUG-III/22 groups and other patient characteristics from the MDS 2.0. The patient-adjusted difference in average nursing time ranged from about -30 to +30 min (Figure 1). Wards at the end of the distribution provided about 30 min more or less wage-adjusted nursing time per patient per day than would be estimated based on patient mix. This indicates that both ward-level effects and unmeasured patient-level effects produce differences in wage-adjusted nursing time between the wards.

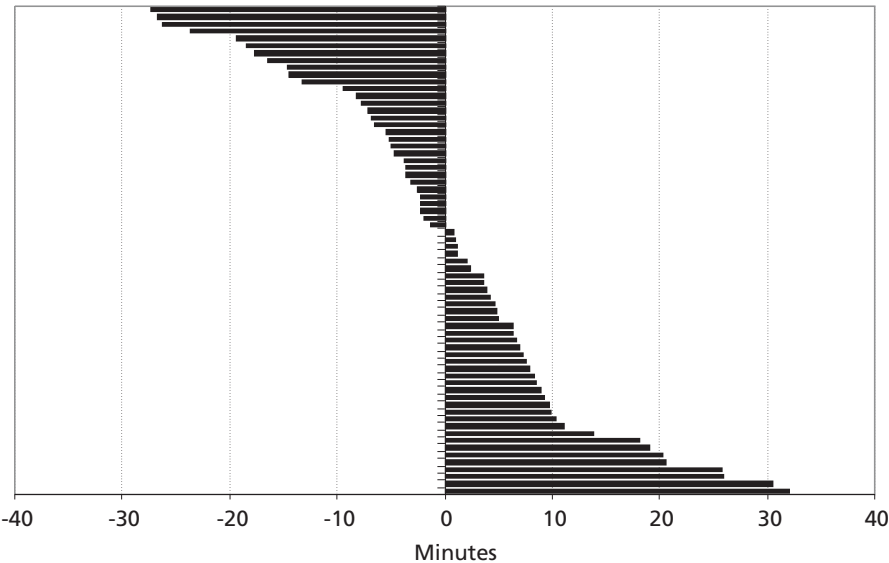


FIGURE 1. The difference between observed and RUG-III/22 predicted wage adjusted nursing time in minutes in the wards (the RUG/22-adjusted model was supplemented with other MDS 2.0 items)

In Table 2, four regression models are presented. The dependent variable in the models is patient’s nursing time in minutes over a 24-h period, which was predicted by patient and ward characteristics. We estimated the OLS model with RUG-III/22 groups, other patient characteristics from the MDS 2.0 and ward dummies (Model 1). In Model 2, RUG-III/22 groups were replaced by our combined classification

of ADL and cognition scales. In Model 3, hierarchical structures of the data were taken into account by applying multilevel modelling, and ward dummies from the first model were replaced by the set of ward characteristics. Multilevel modelling was also applied in Model 4, in which RUG-III/22 groups were replaced by our combined classification of ADL and cognition scales.

TABLE 2. Ordinary least-squares and multilevel regression estimates for models of patient and ward characteristics

Variables	Model 1: RUG-III/22 + ward dummies (OLS) ^{1 2}	Model 2: ADL/CPS + ward dummies (OLS) ¹	Model 3: RUG-III/22 + ward characteristics (multilevel) ²	Model 4: ADL/CPS + ward characteristics (multilevel)
Dependent variable				
Patient's nursing time in minutes over a 24-hour period				
Constant	75 ***	85 ***	43 ***	45 ***
Patient characteristics				
ADL/CPS1	-	34 ***	-	32 ***
ADL/CPS2	-	49 ***	-	48 ***
ADL/CPS3	-	7	-	8
ADL/CPS4	-	27 ***	-	27 ***
ADL/CPS5	-	38 ***	-	37 ***
ADL/CPS6	-	21 ***	-	19 ***
ADL/CPS7	-	27 ***	-	25 ***
ADL/CPS8	-	40 ***	-	38 ***
Age 65 - 74	-6	-5	-8	-6
Age 75 - 84	-12 ***	-10 ***	-14 ***	-12 ***
Age 85 -	-16 ***	-15 ***	-19 ***	-18 ***
Unbalanced relationship with others	-4 ***	-6 ***	-5 ***	-7 **
Daily pain	5 ***	6 ***	6 ***	8 ***
Depression symptoms	10 ***	10 ***	10 ***	11 ***
Admission	-0.003 ***	-0.003 ***	-0.004 ***	-0.003 ***
Nursing rehabilitation	4	8 ***	4 ***	7 ***
Ward characteristics				
Unit size	-	-	0.4	0.3
Shift and team arrangements of the nursing staff	-	-	2.7	4.4
Medications	-	-	0.7	0.5
The proportion of rooms with own bathroom	-	-	0.03	0.04
The proportion of single rooms	-	-	0.06	0.02
The proportion of nursing hours	-	-	0.6 ***	0.7 ***
Ward's specialisation	-	-	1.2	1.1
Facility type	-	-	-7.3	-5.5
R² and variance parameters				
Adjusted R ²	25	19	-	-
Level 1 intercept variance	-	-	1332.46 ***	1450.94 ***
Level 2 intercept variance	-	-	66.60 ***	52.13 ***

* = p<0.05 ** = p<0.01, *** = p<0.001

¹ Ward dummies were used in the model, so ward characteristics were excluded. The parameter estimates of the ward dummies are not presented.

² ADL and CPS variables are part of the RUG-III/22, and therefore included in the RUG-III/22 group dummies. The parameter estimates of the RUG-III/22 groups are not presented.

Both of the OLS models that included patient-level variables were improved by including ward dummies, as they indicated ward-level effects. The adjusted R² was 25 in the RUG-III/22 model and 19 in Model 2. However, the ward characteristics in the multilevel models (Models 3 and 4) were mainly not significant and coefficients were quite low compared to patient characteristics. The proportion of wage-adjusted nursing time out of the total staff working hours was the only ward characteristic that was significant. If the average in the proportion of nursing hours were increased from 49% to 59%, patients would gain 6–7 min of wage-adjusted nursing time per day. In general, the models are very consistent.

The combined measure of the ADL and CPS was the most influential factor

in determining wage-adjusted nursing time in Models 2 and 4, as were the RUG-III/22 dummies in RUG-III Models 1 and 3. The transition from good ADL and good cognition (class 0) to poor ADL and fully impaired cognition (class 8) increased wage-adjusted nursing time by 38 minutes per 24-h according to Model 4. The wage-adjusted nursing times for the different groups of physical functions and cognition are presented in Table 3. The wage-adjusted nursing time per day increased from 68 min at the lowest end of the CPS scale to 87 min at the higher end, if ADL was good or only weakened to a small extent. However, when ADL had deteriorated moderately or severely (ADL 3–6), the level of cognition no longer affected nursing time.

TABLE 3. Nursing minutes on average over a 24-hour period according to physical functions (ADL) and cognition (CPS)¹

	ADL 0-2	ADL 3-4	ADL 5-6
CPS 0-2	68	104	115
CPS 3-4	77	97	107
CPS 5-6	87	94	106

¹ Higher ADL score indicates poorer physical functions, while higher CPS score indicates more impaired cognition. Nursing minutes are wage adjusted.

Other patient-level factors only affected wage-adjusted nursing time to a small extent (Table 2). Patients having symptoms of depression or daily pain were provided with about 10 min more nursing time than patients without daily pain or symptoms of depression. Meanwhile, older patients and patients who had an unbalanced relationship with others were given less nursing time per day compared to younger patients and patients who did not have an unbalanced relationship with others.

Discussion

In Finland the average wage-adjusted nursing time per patient per day has increased from 80 minutes to 100 minutes from 1995 to 2002.¹⁴ Apparently, this is a consequence of both altered nursing patterns and a more complicated patient mix. In this study, the difference between wards in the average wage-adjusted nursing time was marked, which is partly explained by introducing a set of patient-level factors in the analysis. However, the

results indicate that there are differences between wards, even if patient mix is controlled for. The remaining differences in nursing time cannot be explained by the various organizational characteristics. Differences between wards seem to be caused by unmeasured patient and ward level effects, ward level inefficiency and randomness. As a whole, the results indicate an inefficient and inappropriate allocation of the staff, which is consistent with the results of productive efficiency studies in Finnish elderly care.¹³ This shortcoming needs to be given serious consideration by management.

The combined measure of physical functions and cognition provided information on how these factors together affect wage-adjusted nursing time at different levels of impairment. The results also showed that patient characteristics that are not part of the RUG-III case-mix index also affect wage-adjusted nursing time, though only to a quite small extent. Some patient characteristics, such as age over 75 and an unbalanced relationship with others, even affect wage-adjusted nursing time negatively.

The RUG-III/22 explained patients' nursing time to a lesser extent than in previous studies. This may indicate that patients in our time measurement sample ($n = 1797$) were too homogeneous, although the data was gathered from 68 wards. Furthermore, it is possible that poorer quality of care may increase the time patient needs, which was not explored in this study. Hence, further studies with a follow-up sample and risk

adjusted quality measures are warranted to show how much the incidence of poor outcomes and quality affects nursing time.

The different ways of organizing care played a minor role as a determinant of wage-adjusted nursing time in comparison to patient characteristics, unmeasured effects, randomness and inefficiency. Ideally, nursing time should not be determined by extrinsic factors and randomness, but according to each resident's needs for physical help, care, or assurance and comfort. By improving the allocation of labour resources, the marked variation in average wage-adjusted nursing time between wards could be diminished and the efficiency of the use of labour resources improved. Consideration should also be given to the question of whether it is possible to further increase nursing time as a proportion of the total working hours in Finland.

Acknowledgements

This study has been funded by the National Research and Development Centre for Welfare and Health (STAKES), Finland, and the Doctoral Programme in Social and Health Services Management and Economics of the University of Kuopio. We are grateful Suomen Kuntaliitto for supporting the time measurement study. Acknowledgment is also given to interRAI.

References

- Bravo G, De Wals P, Dubois M-F, Charpentier M. Correlates of Care Quality in Long-Term Care Facilities: A Multilevel Analysis. *Journal of Gerontology: Psychological sciences* 1999; 54B: 180-188
- Dellefield ME. The Relationship Between Nurse Staffing in Nursing Homes and Quality Indicators. *Journal of Gerontological Nursing* 2000; 27: 14-28
- Needleman J, Buerhaus P, Mattke S, Stewart M, Zelevinsky K. Nurse-staffing levels and the quality of care in hospitals. *New England Journal of Medicine* 2002; 346: 1715-1722
- GAO. Nursing Homes - Quality of Care More Related to Staffing than Spending. GAO Report, GAO-02-431R. Washington: The General Accounting Office, 2002
- Noro A. Long-Term Institutional Care Among Finnish Elderly Population. Trends and Potential for Discharge. Research Report volume 87. Helsinki: Stakes, 1998
- SVT. Care and Services for Older People 2002. Sosiaaliturva: I. Helsinki: Finnish Official Statistics, 2003; 137
- StatFin. StatFin. In: Finland S, ed volume 2004. Helsinki: Statistics Finland, 2004
- Vaarama M, Luomahaara J, Peiponen A, Voutilainen P. Koko kunta ikääntyneiden asialle. Näkökulmia ikääntyneiden itsenäisen selviytymisen sekä hoidon ja palvelun kehittämiseen. In: Stakes, ed. Raportteja volume 259. Helsinki: Stakes, 2001; 61
- Fries BE, Cooney JR, Leo M. Resource Utilization Groups. A Patient Classification System for Long-term Care. *Medical Care* 1985; 23: 110-122
- Björkgren M. Case-mix Classification and Efficiency Measurement in Long-Term Care of the Elderly. Research Report volume Research report 124. Helsinki: Stakes, 2002
- Fries BE, Schneider DP, Foley WJ, Gavazzi M, Burke R, Cornelius E. Refining a Case-Mix Measure for Nursing Homes: Resource Utilization Groups (RUG-III). *Medical Care* 1995; 32: 668-685
- Arling G, Williams A. Cognitive Impairment and Resource Use of Nursing Home Residents: A Structural Equation Model. *Medical Care* 2003; 41: 802-812
- Laine J, Linna M, Häkkinen U, Noro A. Measuring the Productive Efficiency and Clinical Quality of Institutional Long-Term Care for the Elderly. *Health Economics* 2005; 14: 245-256
- Björkgren M, Häkkinen U, Finne-Soveri H, Fries B. Validity and reliability of Resource Utilization Groups (RUG-III) in Finnish long-term care facilities. *Scandinavian Journal of Public Health* 1999; 27: 228-234
- Takahashi T, Dumbaugh KA, Reich M. Mechanism of Resource Allocation in Long-Term Care. Report. Boston: Harvard School of Public Health, 1995
- Noro A, Finne-Soveri H, Björkgren M, Häkkinen U, Laine J, Vähäkangas P, Kerppilä S, Storbacka R. RAI -tietojärjestelmän käyttöönotto ja pitkäaikaishoidon benchmarking. In: Stakes, ed. RAI -tietojärjestelmän käyttöönotto ja pitkäaikaishoidon benchmarking. Helsinki: Stakes, 2001
- Carpenter I, Perry M. Nursing Hours. Identifying the level of registered nursing care provided to UK nursing home residents: Joseph Rowntree Foundation, 2002; 29
- Morris JN, Hawes C, Fries BE, Phillips CD, Mor V, Katz S, Murphy K, Drugovich ML, Friedlob AS. Designing the National Resident Assessment Instrument for Nursing Homes. *The Gerontologist* 1990; 30: 293-307
- Morris JN, Fries BE, Morris SA. Scaling ADLs within the MDS. *Journal of the Gerontology: Medical Sciences* 1999; 54A: M546-M553
- Morris JN, Fries BE, Mehr DR, Hawes C, Phillips C, Mor V, Lipsitz LA. MDS cognitive performance scale. *Journal of Gerontology: Medical Sciences* 1994; 49: M174-M182
- Fries BE, Simon SE, Morris JN, Flodstrom C, Bookstein FL. Pain in U.S. nursing homes: validating a pain scale for the minimum data set. *Gerontologist* 2001; 41: 173-179
- Burrows A, Morris J, Simon S, Hirdes J, Phillips C. Development of a Minimum Data Set based depression Rating Scale for use in nursing homes. *Age and Ageing* 2000; 29: 165-172

23. Goldstein H. Multilevel Statistical Models: Second Edition Edition. London: Edward Arnold, 1995
24. Rice N, Leyland A. Multilevel models: applications to health data. *Journal of Health Services Research & Policy* 1996; 1: 154-164
25. Rice N, Jones A. Multilevel Models and Health Economics. *Health Economics* 1997; 6: 561-575
26. Rasbash J, Browne W, Goldstein H, Yang M, Plewis I, Healy M, Woodhouse G, Draper D, Langford I, Lewis T. A user's guide to MLwiN. London: Centre for Multilevel Modelling. Institute of Education. University of London, 2001

Tutkimuksia-sarjassa aiemmin ilmestyneet

2005

Manu Kitinoja. Kujan päässä koulukoti. Tutkimus koulukoteihin sijoitettujen lasten lastensuojeluasiakkuudesta ja kouluhistoriasta. Stakes, Tutkimuksia 150. Helsinki 2005. ISBN 951-33-1775-7

2004

Mikko Salasuo. Humeet ajankuvana. Huumeiden viihdekäytön kulttuurinen ilmeneminen Suomessa. Stakes, Tutkimuksia 149. Helsinki 2004. ISBN 951-33-1749-8

Filio Degni. The Social and Cultural Determinants of the Use of Contraception among Married Somali Women Living in Finland. STAKES. Research Report 148. Helsinki 2004. ISBN 951-33-1736-6

Kaija Hänninen. Kohtaamisen kokemuksi epävarmuuden näyttämöllä. Kokemuksellinen ensitieto vammaisen lapsen syntyessä. Stakes, Tutkimuksia 147. Helsinki 2004. ISBN 951-33-1598-3

Susan Kuivalainen. A Comparative Study on Last Resort Social Assistance in Six European Countries. STAKES, Research Report 146. Helsinki 2004. ISBN 951-33-1594-0.

Pasi Moisio. Poverty Dynamics According to Direct, Indirect and Subjective Measures. Modelling Markovian processes in a discrete time and space with error. STAKES, Research Report 145. Helsinki 2004. ISBN 951-33-1587-8

Elina Virokannas. Normaalin rajan molemmilla puolilla. Tutkimus huumehoitoyksikön nuorten identiteettien rakentumisesta. Stakes, Tutkimuksia 144. Helsinki 2004. ISBN 951-33-1586-X

Jenni Leinonen. Families in Stuggle. Child Mental Health and Family Well-being in Finland During the Economic Recession of the 1990s: The Importance of Parenting. STAKES, Research Report 143. Helsinki, Finland 2004. ISBN 951-33-1583-5

Päivi Voutilainen. Hoitotyön laatu ikääntyneiden pitkäaikaisessa laitoshoidossa. Stakes, Tutkimuksia 142. Helsinki 2004. ISBN 951-33-1572-X

Mika Ala-Kauhaluoma, Elsa Keskitalo, Tuija Lindqvist, Antti Parpo. Työttömien aktivointi. Kuntouttava työtoiminta -lain sisältö ja vaikuttavuus. Stakes, Tutkimuksia 141. Helsinki 2004. ISBN 951-33-1571-1

Antti Parpo. Kannustavuutta tulonsiirtojärjestelmään. Tulonsiirtojärjestelmän muutokset, kannustinloukut ja tulonjako. Stakes, Tutkimuksia 140. Helsinki 2004. ISBN 951-33-1568-1

Heikki Hiilamo, Jouko Karjalainen, Mikko Kautto, Antti Parpo. Tavoitteena kannustavampi toimeentulotuki. Tutkimus toimeentulotuen lakimuutoksista. Stakes, Tutkimuksia 139. Helsinki 2004. ISBN 951-33-1567-3

Jukka Kärkkäinen. Onnistuiko psykiatrian yhdentyminen somaattiseen hoitojärjestelmään? Psykiatrisen hoitojärjestelmän kehitys Suomessa ja sairaanhoitopiiriuudistuksen arviointi psykiatrian näkökulmasta. Stakes, Tutkimuksia 138. Helsinki 2004. ISBN 951-33-1565-7

Britta Sohlman. Funtkionaalinen mielenterveyden malli positiivisen mielenterveyden kuvaajana. Stakes, Tutkimuksia 137. Helsinki 2004. ISBN 951-33-1556-8

Liisa Heinämäki: Erityisesti päivähoidossa. Kunnallisten toimijoiden ja päättäjiä näkemykset erityispäivähoidon funktiosta palvelujärjestelmässä. Stakes, Tutkimuksia 136. Helsinki 2004. ISBN 951-33-1555-X

Sirkku Vuorma: A Decision Aid and Choice of Treatment among Women with Heavy Menstrual Bleeding. STAKES. Research Report 135. Helsinki 2004. ISBN 951-33-1554-1

Hannele Hyppönen: Tekniikka kehittyy, kehittyvätkö palvelut? Tapaustutkimus kotipalvelujen kehittämisestä teknologiahankkeessa. Stakes, Tutkimuksia 134. Helsinki 2004. ISBN 951-33-1553-3

Tarja Pösö: Vakavat silmät ja muita kokemuksia koulukodista. Stakes, Tutkimuksia 133. Helsinki 2004. ISBN 951-33-1548-7